Automated Customer Segmentation Based on Smart Meter Data with Temperature and Daylight Sensitivity

Christian Beckel, Leyna Sadamori
Dept. of Computer Science
ETH Zurich, Switzerland
{beckel,sadamori}@inf.ethz.ch

Silvia Santini
Embedded Systems Lab
TU Dresden, Germany
silvia.santini@tu-dresden.de

Thorsten Staake
Energy Efficient Systems
University of Bamberg, Germany
thorsten.staake@uni-bamberg.de

***Abstract*—Utilities increasingly leverage knowledge on their customer's household characteristics in their energy efficiency programs. Examples of such characteristics include the number of persons per household, their employment status, or the type of dwelling they live in. This information allows utilities to personalize energy efficiency campaigns, which increases participation rates and in turn leads to larger energy savings and higher customer retention. However, gathering this information through surveys is costly and cumbersome. We therefore investigate the possibility to automatically infer household characteristics from electricity consumption data measured by an off-the-shelf smart meter. In this paper, we develop a method to determine the sensitivity of a household to outdoor temperature and the times of sunset/sunrise, and use this information to improve the performance of our household classification system. We further investigate the relevance of different features for such a system. Our evaluation is based on smart meter data collected at a 30-minute granularity in more than 4000 Irish households over a period of 75 weeks. The results show that we can improve accuracy by up to 2.3 percentage points using temperature and daylight coefficients. The characteristics floor area, type of dwelling, and percentage of installed energy-efficient light bulbs particularly benefit from temperature and daylight coefficients. Finally, we investigate the impact of the data granularity on the classification performance and show that semi-hourly or hourly data is required, as it performs on average 6.6 percentage points better than using daily consumption averages.**

I. Introduction

Smart electricity meters play an important role in the future electricity grid. They are expected to enable automatic billing, balance demand and supply, and induce energy savings through behavior change by providing consumption feedback to households [1]. For these reasons, more than 50 million smart meters have already been installed in the United States (US), which represents a coverage of 43% nationwide [2]. The European Union (EU) is following suit as it aims at a coverage of 80% until 2020 [3]. At the same time, expectations about the energy savings achievable through the installation of smart meters have diminished. For instance, a meta-analysis of 33 trials involving in-home displays suggests that "3%–5% is a more accurate and applicable expected conservation figure" [4].

When designing energy efficiency programs, utilities must therefore go beyond providing mere consumption feedback to households. In particular the characteristics of a household play an important role when deciding which energy saving measure is most effective to help the household reduce its energy consumption [5]. Households with high income, for instance, are more likely to invest in infrastructural changes, while people aged 65 years and older tend to be more critical towards technical changes and should be targeted by a behavior change program instead [6]. Among the households that are open for infrastructural investments, some might for instance be good targets for a heat pump marketing campaign. Others might instead already have a heat pump installed and are therefore not worth contacting [7]. Similarly, the inhabitants of single- or two-person households typically work during the day. Hence, they most often have a regular schedule and are suitable candidates for a smart thermostat, which automatically controls the heating system of a household depending on its occupancy state [8]. Families, on the other side, should rather be motivated to change their behavior (*e.g.*, reducing room temperature). Ultimately, the effect of consumption feedback is typically higher if advice and motivational cues are tailored to the recipient [9]–[11].

Personalizing energy efficiency programs to individual households therefore benefits from detailed knowledge on the characteristics of the households. However, performing surveys to acquire customer information is costly and cumbersome. For this reason, researchers have investigated the potential of supervised machine learning to automatically infer such information from a household's electricity consumption data [12]–[14]. In this paper, we build on our prior work on automated household classification [12] and present the following contributions:

- We show how to combine smart meter data with outdoor temperature and daylight information to improve the classification performance.
- We present the results of an in-depth analysis of the features used for automated household classification in our system.
- We investigate different smart meter data granularities to learn which temporal granularity is optimal in real-world settings.

Our investigations are performed on the publicly available *CER data set*[1], which contains fine-grained smart meter data collected from several thousand households and small businesses in Ireland. Along with the consumption data, the data set contains answers to more than 100 questionnaire items filled out by 4231 participants, which we use to train the system and evaluate its performance. From those households, we utilize the consumption data recorded in 30-minute intervals from 20 July 2009 to 26 December 2010 (*i.e.*, 75 weeks) in our analysis. The code used to run our analysis is publicly available[2].

II. Automated Household Classification

Our household classification system described in [12] first computes 25 features on each week of the consumption data. Examples of such features are consumption averages during different times of the day, ratios between those averages, statistical features such as cross-correlation or variance, and principal components of the whole feature set. For each week of the data—and for each characteristic—we then

[1] www.ucd.ie/issda/data/commissionforenergyregulationcer/
[2] https://github.com/beckel/class

train a classifier on a dedicated training set. On this training set, we use a feature selection algorithm to automatically identify the set of features that allows our system to achieve the best performance in the classification of household characteristics. With the trained model, we then predict the class of each household in the test set for each week separately and ultimately assign the household to the class that has been predicted for most of the weeks. This process is performed 4-times (*i.e.*, 4-fold cross validation) using four different combinations of the training and test sets such that each household is a part of the test set exactly once. Based on the ground truth information obtained from the questionnaires, we compute the accuracy and Matthews Correlation Coefficient (MCC) [15] to evaluate the performance of our system. We utilize the MCC because accuracy is often considered a weak measure when dealing with imbalanced classes [16]. The MCC ranges between -1 and 1, whereas 1 represents a perfect classification, 0 denotes a classification that is no better than a random classification, and -1 shows a total disagreement between classification and observation.

Our system is able to classify 18 different household characteristics related to the socio-economic properties of the household, properties of the dwelling, or the number and type of appliances. For a detailed description of our system including a definition of all features and characteristics, we refer to our prior work [12].

A. Impact of Temperature and Daylight

Weather information such as the outside temperature has a significant effect on a household's electricity consumption [13], [17], [18]. In this paper, we show that this information can also be used to improve the performance of our system. Since the CER data set does not contain weather information, we extracted temperature readings from online weather data provided by *WeatherOnline*[3]. For each day, WeatherOnline provides weather parameters at a 30-minute granularity. These include wind speed, temperature, humidity/visibility, precipitation, clouds, and pressure. WeatherOnline's data has been collected from different weather stations across Ireland, whereas not all stations provide data at all times. Since the CER data set does not provide information on the location of the dwellings of the participating households, we assume that all households experience the same outdoor temperature. To this end, we compute an average value of three weather stations across the country that provided information for almost all the 30-minute time slots during the study. These stations are: Cork-Corcaigh, located in the south of Ireland at 162 m altitude; Dublin Airport, located in the East of Ireland at 85 m altitude; and Shannon, located in the West of the country at 20 m altitude. The distance between the three stations is only approximately 200 km. From these three weather stations we extracted semi-hourly temperature readings and used their average for the rest of the analysis.

Figure 1 shows the correlation between electricity consumption and outside temperature. Each of the 525 points on the plot corresponds to on the y-axis the average electricity consumption of all 4231 households on a particular day and on the x-axis the average outside temperature of that day. The plot shows that the electricity consumption increases as the outside temperature decreases. This is in part due to the fact that heating systems consume more electricity when it is cold outside. However, for the households in the CER data set, most of the dwellings are heated using fossil fuels, while only a small percentage of households uses a central electric heating system or an electric plug-in heater. Besides

[3]www.weatheronline.co.uk/weather/maps/current?CONT=euro&LAND=IE

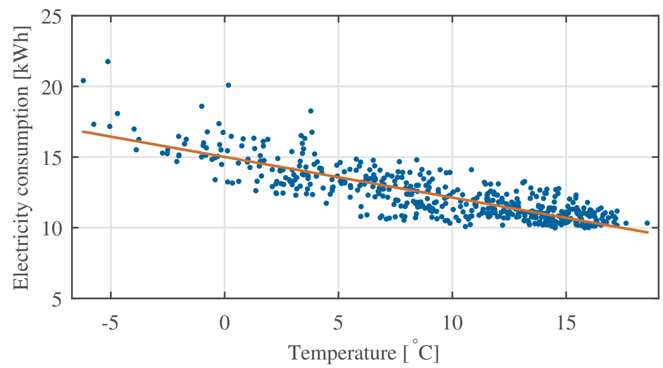

Figure 1: Correlation between average daily electricity consumption of all households and outside temperature.

electricity consumed by space heating, lifestyle also contributes to the increase in electricity consumption when temperature decreases. For instance, people spend more time at home—and thus consume more electricity—when it is cold outside.

In addition to the sensitivity to outside temperature, we investigate whether the correlation between electricity consumption and daylight is a useful feature to infer household characteristics. To this end, we extracted the times of sunrise and sunset for each day using the Astronomy API from *timeanddate.com*. To extract the sensitivity of each household to outside temperature and daylight, we use the linear regression model specified by

$$y_{jt} = \alpha_j + \sum_{k=1}^{3} \beta_j^{(k)} W_t + \sum_{k=1}^{24} \gamma_{jk} \mathbb{1}\{\mathrm{ToD}(t) = k\} + \sum_{k=1}^{2} \delta_{jk} \mathbb{1}\{\mathrm{Weekday}(t) = k\} + \varepsilon_{jt}, \quad (1)$$

where y_{jt} represents the j-th household's electricity consumption at time $t \in T$, $T = 1, ..., 25200$. T thus represents all 30-minute time slots during the 75 weeks of the trial. The coefficient α_j denotes a constant term for each household. γ and δ are dummy variables for time of day (in hours) and for the fact whether t is on a weekday or on the weekend, respectively. The coefficient ε represents the error term.

The three coefficients $\beta^{(1)}$, $\beta^{(2)}$, and $\beta^{(3)}$ account for the sensitivity to sunrise, sunset, and temperature, respectively. These values are incorporated in W_t, given by

$$W_t = \begin{pmatrix} \mathrm{Sunrise}(t) \\ \mathrm{Sunset}(t) \\ \mathrm{Temperature}(t) \end{pmatrix}, t \in T. \quad (2)$$

Temperature(t) denotes the outdoor temperature at time t, which we compute on the data collected by the weather stations. Further, we determine in which time slot of the day sunrise and sunset occur. Using the number of minutes between midnight and sunrise (`min_sunrise`) and the number of minutes between midnight and sunset (`min_sunset`) from the astrology data, we compute

$$\texttt{sunrise\_slot} = \texttt{min\_sunrise}/30 \quad (3)$$

and

$$\texttt{sunset\_slot} = \texttt{min\_sunset}/30. \quad (4)$$

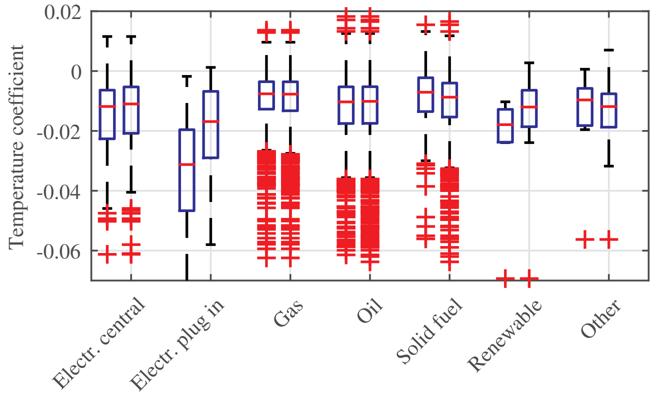

Figure 2: Temperature coefficients obtained from the regression analysis categorized by type of space heating. For each type, the left side shows the coefficients of the households with only a single heat source, and the right side shows the coefficients of all households.

We then set Sunrise(t) and Sunset(t) to 1 if t lies within the hour that follows sunrise or within the hour that precedes sunset, respectively:

$$\text{Sunrise}(t) = \begin{cases} 1 & \text{if } t \pmod{\lceil \texttt{sunrise\_slot} \rceil + 1} = 1 \\ & \quad \vee t \pmod{\lceil \texttt{sunrise\_slot} \rceil + 2} = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

$$\text{Sunset}(t) = \begin{cases} 1 & \text{if } t \pmod{\lceil \texttt{sunset\_slot} \rceil} = 1 \\ & \quad \vee t \pmod{\lceil \texttt{sunset\_slot} \rceil - 1} = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

For instance, if sunrise is at 5:15 a.m. and sunset at 9:15 p.m., the 12th and the 13th time slots of the day indicate sunrise (*i.e.*, from 5:30 a.m. to 6:30 a.m.), and the 42th and 43rd time slots of the day (*i.e.*, from 8:30 p.m. to 9:30 p.m.) indicate sunset. By introducing the dummy variables for time of day and weekday/weekend we ensure that similar time slots are used to estimate the effect of sunrise, sunset, and outdoor temperature on electricity consumption data.

Figure 2 shows a box plot of the temperature coefficients $\beta^{(3)}$ for each household obtained from the regression analysis. For each heat source shown on the x-axis, the left side includes those households that specified only a single heat source in the questionnaire. The plot shows the strongest correlation between electricity consumption and outdoor temperature for those households that have an electric plug-in heater installed in their household. The second strongest correlation is shown for those households that use renewables to heat their home. A potential reason for this is that those households generate most electricity during warm, summer days. The effect for households that heat with gas, oil, or solid fuels is also visible, although lower than for those households with electric heating. Most likely the sensitivity to temperature for those households that heat with fossil fuels stems from the fact that the lifestyle of inhabitants depends on the outdoor temperature. The right side of each heat source in the plot shows the temperature coefficients for all households including those that have multiple heat sources according to the answers to the questionnaires. The effects described above are still visible. However, they are much lower due to the fact that the data set does not specify which of the heat sources is mainly used in a household. For instance, a household that mostly heats by gas but also owns a (rarely used) plug-in heater contributes to both the second and the third column of the plot, evening out effects of households that only use a plug-in heater. In fact, only 108 households do not heat their house using oil, gas, or solid fuels, and from those households, 76 households rely on central electric heating. Only 16 households heat their house using a plug-in heater only. Overall, the analysis shows that energy consumed by space heating—at least as the primary heat source—is barely accounted for in the consumption data of the CER data set.

To evaluate the use of weather coefficients for household classification, we include $\beta^{(1)}$, $\beta^{(2)}$, and $\beta^{(3)}$ as additional input features. In particular, we run the classification two times for each of the 75 weeks in the data set using the Linear Discriminant Analysis (LDA) classifier, which is the fastest classifier supported by our system and provides similar results to other classifiers [12]. In the first run, we use accuracy as figure of merit during feature selection. The second time we run the analysis using MCC as figure of merit instead. Table I shows the results of the analysis and compares it to the results obtained in [12]. The left side of the table shows the accuracy obtained for each of the characteristics. Adding daylight and temperature coefficients to the feature set improved accuracy for characteristics `floor_area` (increase by 2.1 percentage points from 50.5% to 52.6%), `house_type` (increase by 2.3 percentage points from 59.3% to 61.5%), and `lightbulbs` (increase by 1.1 percentage points from 55.1% to 56.2%). The other characteristics exhibit a negligible increase, decrease, or no change. A decrease can occur in cases where the new feature set performs better than the previous features on the training data but worse on the test data. Overall, accuracy increased by 0.29 percentage points from 65.1% to 65.4%. In terms of MCC, we observe a large increase for characteristics `house_type` (increase by 0.046 points from 0.191 to 0.236), `lightbulbs` (increase by 0.018 from 0.0976 to 0.116), and `#appliances` (increase by 0.011 from 0.343 to 0.354), achieving an overall increase by 0.0036 points from 0.317 to 0.32.

Overall, the weather coefficients (*i.e.*, daylight and temperature features) increase performance for characteristics related to the dwelling (*i.e.*, `floor_area` and `house_type`) or to the efficiency of lighting (*i.e.*, `lightbulbs`). While the former increase is possibly caused by the temperature coefficient, the latter can most likely be attributed to the inclusion of sunrise and sunset into the analysis. In the next section, we present a detailed analysis on which features are selected by the feature selection algorithm for different household characteristics.

B. Features

In this subsection, we evaluate the importance of each individual feature used in our analysis. To this end, we exclude the principal components from the feature set and run the analysis including the daylight and temperature coefficients. The feature selection algorithm described in [12] selects a maximum of 8 features per run, and it only adds a feature to the feature set if the figure of merit improves significantly (*i.e.*, by at least 0.005). Since we perform 4-fold cross validation, this results in up to 32 selected features for each characteristic, whereas each individual feature can be selected up to 4 times. Overall, since we run the classification separately on each of the 75 week's data, each feature can be selected up to 300 times for each characteristic, or up to 5400 times in total.

Figure 3 shows the number of times each feature was selected. The colors of the bars indicate the type of the feature: The first 10

Table I: Results obtained by classifying 4231 households using the original features [12] and the new features including daylight and temperature coefficients ("+ Weather"). Column 3 shows the classes for each characteristic. A detailed specification of each class is provided in [12].

			Accuracy			MCC		
Characteristic	**Description**	**Classes**	**Original**	**+ Weather**	**Change**	**Original**	**+ Weather**	**Change**
`single`	Single	{*Single, No single*}	82%	81.8%	-0.15	0.495	0.503	+0.0077
`all_employed`	All adults work for pay	{*Yes, No*}	78.6%	78.6%	0	0.324	0.313	-0.011
`unoccupied`	Is the house unoccupied for more than 6 hours per day?	{*Yes, No*}	76.4%	76.4%	-0.025	0.376	0.38	+0.0039
`family`	Family	{*Family, No family*}	73.7%	73.5%	-0.22	0.419	0.415	-0.0043
`children`	Children	{*Yes, No*}	72.8%	72.5%	-0.35	0.408	0.407	-0.0015
`cooking`	Type of cooking facility	{*Electrical, Not electr.*}	71.2%	71.2%	0	0.286	0.29	+0.0041
`retirement`	Retirement status of chief income earner (CIE)	{*Retired, Not retired*}	73.5%	73.8%	0	0.43	0.427	-0.0036
`#residents`	Number of residents	{*Few, Many*}	75.5%	75.3%	-0.2	0.508	0.513	+0.0045
`employment`	Employment of CIE	{*Employed, Not empl.*}	72.3%	72.2%	-0.12	0.436	0.443	+0.0079
`floor_area`	Floor area	{*Small, Medium, Big*}	50.5%	52.6%	**+2.1**	0.205	0.197	-0.0085
`age_person`	Age of CIE	{*Young, Medium, High*}	58.6%	59.3%	0	0.295	0.297	+0.0018
`age_house`	Age of building	{*Old, New*}	63.7%	63.9%	0	0.278	0.277	-0.0014
`house_type`	Type of house	{*Free, Connected*}	59.3%	61.5%	**+2.3**	0.191	0.236	**+0.046**
`income`	Yearly household income	{*Low, High*}	61.1%	60.5%	-0.54	0.229	0.214	-0.015
`lightbulbs`	Proportion of energy-efficient light bulbs	{*Few, Many*}	55.1%	56.2%	**+1.1**	0.0976	0.116	**+0.018**
`social_class`	Social class of CIE	{*A/B, C1/C2, D/E*}	52.9%	52.9%	0	0.225	0.227	+0.0016
`#appliances`	Number of appliances	{*Low, Medium, High*}	55.7%	56.2%	0	0.343	0.354	**+0.011**
`#bedrooms`	Number of bedrooms	{*1–2, 3, 4, >4*}	38.7%	38.4%	-0.27	0.151	0.155	+0.0039
Mean			**65.1%**	**65.4%**	**+0.29**	**0.317**	**0.32**	**+0.0036**

features (blue bars) represent consumption figures, the next 7 features (orange bars) consumption ratios, the next 4 features (yellow bars) temporal properties, the next 4 features (purple bars) statistical properties and ultimately, the last 3 features (green bars) represent the weather coefficients introduced in this paper. For space reasons, we only present the results obtained using accuracy as figure of merit. The results obtained using MCC are very similar. The black, dotted line indicates the median value. The plot shows that the features that represent average consumption data are roughly equally distributed. Exceptions are features `c_evening` and `c_min`, which have been selected 1214 and 1261 times, respectively, and thus more often than the others. Similarly, the features that represent consumption ratios are roughly equally distributed, except `r_evening/noon`, which was selected 1587 times. The feature that indicates if a household's consumption reaches 0.5 kW throughout a day was selected most often: in 2469 out of the 5400 classifications, `t_above_0.5kw` was selected by the feature selection algorithm. Ultimately, weather coefficients, in particular `w_sunset` and `w_temperature`, have been selected more often than many other features with 1552 and 1618 selections, respectively.

Overall, the features selected most often are two different consumption features representing the average and minimum consumption, a consumption ratio which divides the evening consumption by the noon consumption, the time period in which a household's consumption is above 0.5 kW, and weather coefficients (sunset and temperature). The statistical properties have been rarely selected by the feature selection algorithm.

In figure 4, we show the number of selections per feature for each household characteristic. The plot shows—for each feature and household characteristic—the number of times the particular feature has been selected indicated by the color of the particular segment. The total average consumption of a household (`c_total`), for instance, plays a major role when classifying the characteristics related to the number of persons or appliances in the household (*i.e.*, `#children`, `#residents`, `family`, and `#appliances`). Consumption ratios are often included into the feature set when classifying characteristics related to the occupancy of the house (*i.e.*, `employment`, `retirement`). Finally, weather-related features play an important role when classifying characteristics related to the number of appliances (*i.e.*, `#appliances`) or the dwelling (*i.e.*, `#bedrooms`, `floor_area`, `house_type`). For those characteristics with a relatively uneven distribution of samples per class (*i.e.*, `age_person`, `all_employed`, `single`, and `unoccupied`), using accuracy as figure of merit leads to the selection of only a few distinct features.

C. Data Granularity

To evaluate the impact of smart meter data granularity on the performance of our household classification system, we transformed the data to obtain *30-minute data* (no averages), *60-minute data* (averages of two consecutive readings) and *daily data* (averages of 48 consecutive readings). For both 60-minute data and daily data we adapted the features accordingly. The former case only requires recomputing the existing features on the new data. For the daily data, however, many of the features such as the consumption ratios (except `r_weekday/weekend`) and statistical features (*e.g.*, variance and cross-correlation) must be omitted. To compute the sensitivity of each household to outdoor temperature and daylight given 60-minute data, we performed the regression analysis as described before except that we also aggregated the temperature readings accordingly. In the case of daily data, we omitted the daylight features and computed temperature sensitivity based on daily electricity consumption and temperature data. Using these modified feature sets, we performed the analysis as described above, using 75 weeks of data and both accuracy and MCC independently as figure of merit during feature selection.

Figure 5 shows the results obtained when using accuracy as figure of merit. Each set of bars illustrates the accuracy obtained when running the classification on 30-minute data (left bar), 60-minute data (center bar), and daily data (right bar). The plot shows that the difference between 30-minute data and 60-minute data is negligible. On average, accuracy is 0.3

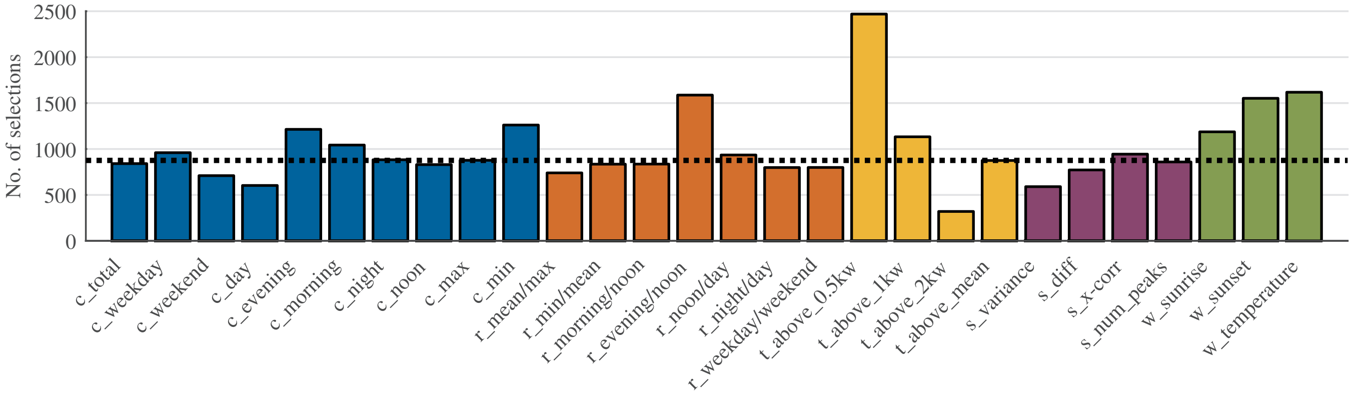

Figure 3: Number of times each feature has been selected by the feature selection algorithm (figure of merit: Accuracy).

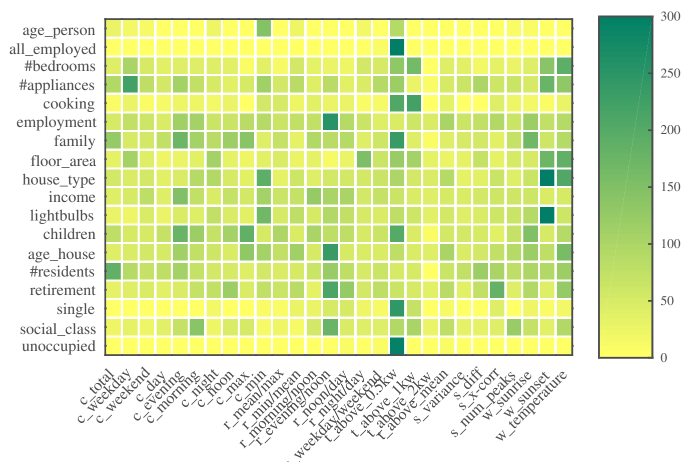

Figure 4: Number of times each feature has been selected by the feature selection algorithm for each characteristic (figure of merit: accuracy).

percentage points higher when 30-minute data is available. However, for some of the characteristics, classification using 60-minute data shows slightly better results. While this seems counterintuitive, a possible explanation is that different features are selected during training, which can result in a better performance on the test set. Comparing the results obtained on 30-minute data with the results obtained when classifying on daily data shows that the accuracy using 30-minute data is much higher (*i.e.*, 6.6 percentage points on average). This is in particular visible for characteristics that make use of consumption ratios: Classifying characteristics `all_employed`, `unoccupied`, `retirement`, and `employment` shows a difference of 14.9, 14.1, 12.0, and 11.8 percentage points, respectively. Characteristics `#appliances` and `#residents`, however, correlate stronger with the overall electricity consumption, which is indicated by a relatively low difference between 30-minute data and daily data with 1.3 and 1.4 percentage points, respectively. Overall, the analysis shows that in practical settings, capturing data at daily granularity is not sufficient, while the additional overhead in collecting 30-minute data instead of 60-minute data does not pay off.

III. Related Work

A popular line of research in smart meter data analytics consists of disaggregating the consumption of individual appliances from a household's aggregate electricity consumption [19]. This approach called non-intrusive load monitoring (NILM) allows to provide automated energy-saving recommendations and is considered the "holy grail of energy efficiency" [20]. However, NILM requires consumption data in the order of 1 Hz or multiple kilohertz.

Approaches that focus on the analysis of more coarse-grained electricity consumption data often aim at different types of customer segmentation: Kwac *et al.* for instance analyze the lifestyle of customers based on their electricity consumption in order to optimize demand side management programs [21]. In [13], Albert *et al.* perform an approach similar to ours as they infer household characteristics from smart meter data. In contrast to the approach described in this paper, the authors even out the effect of weather on the consumption data and compute features on the residuals. Also, the data set used to evaluate their approach is significantly smaller than ours. Other approaches that investigate the correlation between electricity consumption and household characteristics are performed by Kavousian *et al.* [17], [22] and by McLoughlin *et al.* [23]. However, in contrast to our approach, they do not aim at inferring household characteristics from consumption data.

An important aspect in the context of smart meter data analytics is privacy. In general, it is important to inform the users on what happens to their data and find a balance between privacy concerns and legitimate applications [24]. Other approaches even aim at hindering applications such as the ones provided in this paper, for instance by using a battery to "obfuscate" a household's consumption data [25].

IV. Conclusion

With the advent of smart metering, inferring household characteristics from electricity consumption data becomes an important tool for utilities to offer personalized energy efficiency programs at large scale. In this paper, we show how to improve household classification by including outdoor temperature and daylight coefficients into the analysis. The results show a performance increase for certain characteristics. This pays off in practice provided that the number of customers is large and these parameters can be extracted for free from public resources. We further provide an in-depth feature analysis, in which we identify the features that are particularly relevant for our household classification system. Finally, we show that in order to infer household characteristics from smart meter data, collecting 30-minute data or 60-minute data is a necessity. In our study, performance on such data is on average 6.6 percentage points better than the results obtained when inferring household characteristics from daily data.

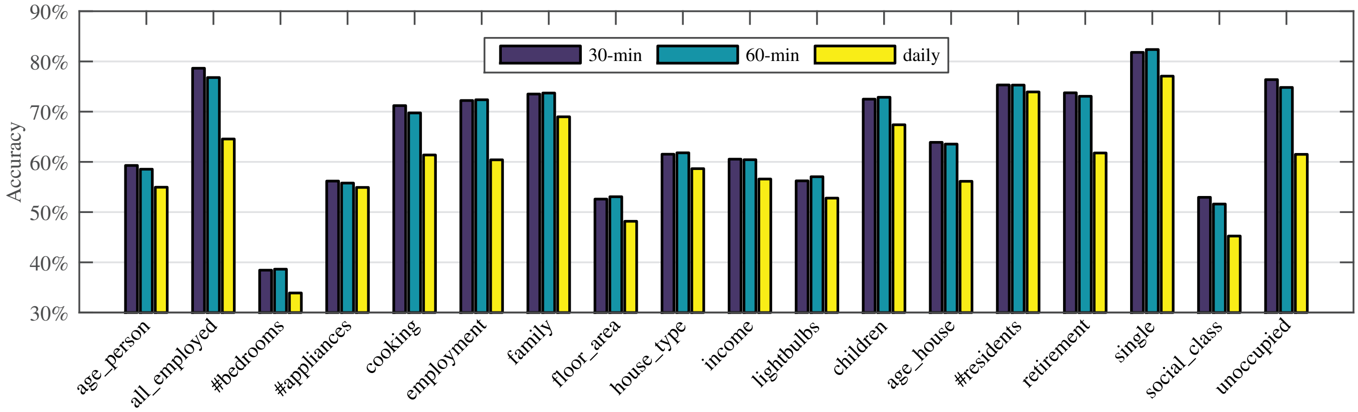

Figure 5: Accuracy for each characteristic achieved when running the analysis on different types (*i.e.*, granularities) of smart meter data.

Acknowledgments

This work has been supported by the Swiss Federal Office of Energy within the program "Electricity Technology and Applications" and by the Hans L. Merkle Foundation. We also thank Prof. Friedemann Mattern and Wilhelm Kleiminger for their valuable comments and for the fruitful discussions about the topic. Finally, we want to express our gratitude to our anonymous reviewers for their remarks.

References

[1] H.-J. Appelrath, H. Kagermann, and C. Mayer (Eds.), "Future energy grid. Migration to the Internet of energy. acatech STUDY," Oct. 2012.

[2] The Edison Foundation, "Utility-scale smart meter deployments: Building block of the evolving power grid," Sep. 2014.

[3] Directive 2009/72/EC of the European Parliament and of the Council of 13 July 2009 concerning common rules for the internal market in electricity and repealing Directive 2003/54/EC.

[4] C. McKerracher and J. Torriti, "Energy consumption feedback in perspective: Integrating Australian data to meta-analyses on in-home displays," *Energy Efficiency*, vol. 6, no. 2, pp. 387–405, May 2013.

[5] C. Beckel, L. Sadamori, and S. Santini, "Towards automatic classification of private households using electricity consumption data," in *4th ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (BuildSys)*. ACM, Nov. 2012, pp. 169–176.

[6] W. Poortinga, L. Steg, C. Vlek, and G. Wiersma, "Household preferences for energy-saving measures: A conjoint analysis," *Journal of Economic Psychology*, vol. 24, no. 1, pp. 49–64, Feb. 2003.

[7] H. Fei, Y. Kim, S. Sahu, and M. Naphade, "Heat pump detection from coarse grained smart meter data with positive and unlabeled learning," in *19th ACM International Conference on Knowledge Discovery and Data Mining (KDD)*. ACM, Aug. 2013, pp. 1330–1338.

[8] W. Kleiminger, F. Mattern, and S. Santini, "Predicting household occupancy for smart heating control: A comparative performance analysis of state-of-the-art approaches," *Energy and Buildings*, vol. 85, pp. 493–505, Dec. 2014.

[9] H. Allcott, "Social norms and energy conservation," *Journal of Public Economics*, vol. 95, no. 9, pp. 1082–1095, Oct. 2011.

[10] I. Ayres, S. Raseman, and A. Shih, "Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage," *Journal of Law, Economics, and Organization*, vol. 29, pp. 992–1022, Oct. 2013.

[11] N. J. Goldstein, R. B. Cialdini, and V. Griskevicius, "A room with a viewpoint: Using social norms to motivate environmental conservation in hotels," *Journal of Consumer Research*, vol. 35, no. 3, pp. 472–482, Mar. 2008.

[12] C. Beckel, L. Sadamori, T. Staake, and S. Santini, "Revealing household characteristics from smart meter data," *Energy*, vol. 78, pp. 397–410, Dec. 2014.

[13] A. Albert and R. Rajagopal, "Smart meter driven segmentation: What your consumption says about you," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4019–4030, Jun. 2013.

[14] K. Hopf, M. Sodenkamp, I. Kozlovkiy, and T. Staake, "Feature extraction and filtering for household classification based on smart electricity meter data," *Computer Science – Research and Development*, pp. 1–8, Nov. 2014.

[15] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochimica et Biophysica Acta (BBA)-Protein Structure*, vol. 405, no. 2, pp. 442–451, Oct. 1975.

[16] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, Jul. 2009.

[17] A. Kavousian, R. Rajagopal, and M. Fischer, "Ranking appliance energy efficiency in households: Utilizing smart meter data and energy efficiency frontiers to estimate and identify the determinants of appliance energy efficiency in residential buildings," *Energy and Buildings*, vol. 99, pp. 220–230, Jul. 2015.

[18] B. J. Birt, G. R. Newsham, I. Beausoleil-Morrison, M. M. Armstrong, N. Saldanha, and I. H. Rowlands, "Disaggregating categories of electrical energy end-use from whole-house hourly data," *Energy and Buildings*, vol. 50, pp. 93–102, Jul. 2012.

[19] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," in *29th International Conference on Consumer Electronics (ICCE)*. IEEE, Januaray 2011, pp. 239–240.

[20] K. C. 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, Jan. 2013.

[21] J. Kwac, C.-W. Tan, N. Sintov, J. Flora, and R. Rajagopal, "Utility customer segmentation based on smart meter data: Empirical study," in *4th International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, Oct. 2013, pp. 720–725.

[22] A. Kavousian, R. Rajagopal, and M. Fischer, "Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior," *Energy*, vol. 55, pp. 184–194, May 2013.

[23] F. McLoughlin, A. Duffy, and M. Conlon, "Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: An Irish case study," *Energy and Buildings*, vol. 48, pp. 240–248, May 2012.

[24] E. McKenna, I. Richardson, and M. Thomson, "Smart meter data: Balancing consumer privacy concerns with legitimate applications," *Energy Policy*, vol. 41, pp. 807–814, Feb. 2012.

[25] D. Egarter, C. Prokop, and W. Elmenreich, "Load hiding of household's power demand," in *5th International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, Nov. 2014, pp. 854–859.