

# Automatic Socio-Economic Classification of Households Using Electricity Consumption Data

Christian Beckel<sup>\*</sup>, Leyna Sadamori<sup>\*†</sup>, Silvia Santini<sup>†</sup>

<sup>\*</sup>Institute for Pervasive Computing, ETH Zurich, Switzerland

<sup>†</sup>Wireless Sensor Networks Lab, TU Darmstadt, Germany

{beckel,sadamori}@inf.ethz.ch, santinis@wsn.tu-darmstadt.de

## ABSTRACT

Interest in analyzing electricity consumption data of private households has grown steadily in the last years. Several authors have for instance focused on identifying groups of households with similar consumption patterns or on providing feedback to consumers in order to motivate them to reduce their energy consumption. In this paper, we propose to use electricity consumption data to classify households according to pre-defined “properties” of interest. Examples of these properties include the floor area of a household or the number of its occupants. Energy providers can leverage knowledge of such household properties to shape premium services (e.g., energy consulting) for their customers. We present a classification system – called CLASS – that takes as input electricity consumption data of a private household and provides as output the estimated values of its properties. We describe the design and implementation of CLASS and evaluate its performance. To this end, we rely on electricity consumption traces from 3,488 private households, collected at a 30-minute granularity and for a period of more than 1.5 years. Our evaluation shows that CLASS – relying on electricity consumption data only – can estimate the majority of the considered household properties with more than 70% accuracy. For some of the properties, CLASS’s accuracy exceeds 80%. Furthermore, we show that for selected properties the use of a priori information can increase classification accuracy by up to 11%.

## Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

## Keywords

Energy consumption analysis; smart electricity meters; household classification; machine learning;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*e-Energy’13*, May 21–24, 2013, Berkeley, California, USA.

Copyright 2013 ACM 978-1-4503-2052-8/13/05 ...\$15.00.

## 1. INTRODUCTION

Electricity providers typically need to collect, process, and store consumption data of their customers for billing purposes. In the past, such consumption data has been collected at a very coarse granularity (e.g., monthly or even yearly). More recently, technological developments as well as changing market rules and regulatory frameworks have made the collection of more fine-grained data (e.g., hourly readings) both an opportunity and a necessity [1, 2, 3]. Electricity providers in Europe and elsewhere are thus putting in place new – or adapting existing – infrastructures to collect, process, and store the expected large amounts of electricity consumption data [1]. Beyond enabling billing according to dynamic pricing policies [3], the availability of such fine-grained data also opens up other opportunities. Several authors have for instance focused on using this data to provide feedback to customers about their electricity consumption [4, 5, 6]. Some studies have also looked at the possibility of deriving the consumption of single household appliances by analyzing the aggregated consumption curve of a private household [7, 8, 9]. Others have focused on discovering consumption curves with similar temporal patterns in order to better tune load prediction algorithms [10, 11].

In this paper, we argue that through the analysis of electricity consumption data it is possible to infer with high probability specific characteristics of private households – like their floor area or the number of persons living in them. We refer to these characteristics as the *properties* of a household. In a previous study, we have defined a set of properties that can be considered both useful to know as well as likely to be inferable with reasonable accuracy from electricity consumption data [12]. In this paper, we build upon our previous work and present the design, implementation, and evaluation of CLASS – a system that estimates the value of the properties of a household for which electricity consumption data is available.

The possibility to derive information about household properties in an automated fashion enables the development of novel services for electricity consumers. Indeed, the ongoing liberalization of electricity markets worldwide makes it easier for consumers to change their providers. These, in turn, need to increasingly offer premium services in order to attract new customers – as well as to retain existing ones. Through in-depth interviews with employees of four different Swiss energy providers, we have identified *energy consulting* as a representative example of such services [12]. The goal of

energy consulting is to provide practical recommendations to customers in order to allow them to reduce their overall energy consumption, thus saving money. By providing such services, providers can not only retain existing customers or attract new ones, but also comply with the ever increasing societal and political pressure to reduce energy consumption in general.

Currently, the development of energy consulting is hampered by the difficulty of identifying customers who might both be interested in and profit from a consulting session. Furthermore, energy consulting sessions are still expensive as they typically involve a personal visit of a consultant to the customers' home. Our interviews have outlined that both the choice of target customers as well as the preparation – and thus effectiveness – of energy consulting sessions would be significantly improved if information about the properties of the customer's household was available in advance. For instance, households occupied by elderly people are more likely to be interested in – and be receptive to – consultants' recommendations [12]. Information about household properties is however typically not available, as providers know surprisingly little about their customers. This is particularly true in European countries in which open information repositories like public tax registers do not exist. Given this restriction, such information has to be acquired through expensive and time-consuming customer surveys. CLASS addresses this problem by providing a way to automatically estimate the properties of private households using electricity consumption data only – which energy providers already have for the purpose of billing. In this way, CLASS enables the development of energy consulting as well as of other premium services. In our work, we implicitly assume that providers will handle electricity consumption data according to existing regulations and industry practices with respect to the processing and propagation of personal data (e.g., [13, 14]).

We formulate the problem of determining the value of a property for a specific household as a classification problem. For a property that can take  $K$  different values, the household is accordingly assigned to one of  $K$  classes. CLASS takes as input a set of features computed over the consumption data of a household and returns as output – for each property – the class to which the household belongs. For instance, in our design the property `floor_area` can take one of three values: small ( $\leq 100 m^2$ ), medium ( $> 100 m^2$  and  $\leq 200 m^2$ ), or large ( $> 200 m^2$ ). CLASS can thus classify a household as a member of one of the three classes *small*, *medium* or *large*, thereby providing an estimation of the value of the corresponding property.

We have analyzed the performance of our classification system for twelve different properties and using four different classifiers. We have trained and evaluated the system using electricity consumption data of 3,488 households. The data traces have been collected over a period of about 1.5 years and at a granularity of 30 minutes (i.e., a data sample represents the average electricity consumption in kilowatts during a 30-minute interval). For all used data traces, we know the actual values of all properties for which the classification system is evaluated. This information – along with the consumption data itself – has been gathered in the context of a smart metering study conducted by the Irish Commission for Energy Regulation (CER). During this study, the electricity consumption data of 4,225 private households and 485 small and medium enterprises has been collected.

All participants of the study have also provided information about themselves by filling in detailed questionnaires. The twelve household properties used for the evaluation of our system have been extracted from these questionnaires as described in [12]. Both the electricity consumption data as well as the questionnaires – to which we refer to as the *CER data set* – have been recently made available to the public.<sup>1</sup>

Our results show that CLASS makes automatic classification of household properties feasible and reliable. In particular, CLASS's classification accuracy is higher than 70% for eight of the twelve properties considered in this study. For two properties, the accuracy exceeds 80%. Furthermore, the use of a priori information allows to further improve CLASS's performance by up to 11%. To the best of our knowledge, this is the first study that provides a quantitative analysis of the possibility to extract households' properties from electricity consumption data. This is also due to the fact that large, labeled data sets of electricity consumption data have only very recently been made available to the public.

The remainder of this paper is organized as follows. We review related work in section 2 and then present the design of CLASS in section 3. In sections 4 and 5 we present the setup and results of our evaluation. We discuss the limitations of our approach and interesting directions for future work in section 6.

## 2. RELATED WORK

Measuring and analyzing fine-grained electricity consumption data has gained attention of many researchers. One of the applications often referred to is non-intrusive load monitoring (NILM) [7, 8, 9, 15, 16]. NILM researchers aim at deriving detailed consumption information – such as the time period when certain appliances are running or how much electricity they consume – from the aggregated electricity consumption of the household. In [7], Zeifman et al. provide a review of established techniques, highlighting different types of features that can be extracted from the consumption data, as well as requirements to the data itself. With such detailed information available, behavioral studies increasingly investigate its use to motivate a more thrifty use of energy [5, 6]. Reinhardt et al. perform a different approach as they detect which appliances are currently running based on consumption data measured at individual power outlets [16]. In practice, NILM approaches suffer from the requirements to high frequency consumption measurements [8] – typically in the order of 1 Hz to multiple kilohertz – as well as from intensive training procedures [7]. Our work is substantially different to NILM as we do not aim at determining the consumption of individual appliances but at identifying high-level characteristics of a household – like the floor area or the number of people living in it. For this purpose, data with a much coarser granularity (e.g., one measurement every 30 minutes) is sufficient, although – in contrast to NILM – such a study requires a comparison of hundreds or thousands of households.

Focusing on data measured with a granularity of 15, 30, or 60 minutes, other related approaches detect consumption patterns of households over a long time frame [11, 17, 18, 19]. To energy providers, applications based on such data are of particular interest as this is the type of data that was col-

<sup>1</sup>[www.ucd.ie/issda/data/commissionforenergyregulation/](http://www.ucd.ie/issda/data/commissionforenergyregulation/)

lected during most of the smart meter trials so far. To identify patterns in electricity consumption data, for instance, De Silva et al. propose a data mining framework and introduce an incremental learning algorithm that predicts future electricity usage of private households [11]. Abreu et al. similarly focus on supporting the energy provider’s supply management as they employ pattern recognition techniques to recognize specific consumption patterns such as daily routines or consumption baselines [17]. By clustering households based on their electricity consumption, Verdu et al. can recognize consumers with an “atypical” consumption pattern [18]. The authors employ a clustering technique – based on self-organizing maps – to classify (commercial) customers by their type. The results presented in these papers are based on relatively small data sets containing traces from 5–30 households. In [19], Birt et al. disaggregate the total electricity consumption into load categories. Evaluating hourly consumption data from 327 households, the authors observe correlations between the consumption of a household and the air temperature both in winter and summer. These correlations allow to determine the electricity consumed by heating and cooling systems, respectively. Our work differs from all approaches above in that we detect properties of households rather than consumption patterns or usage categories. We further rely on a significantly larger data set as we analyze consumption traces from 3,488 households.

Many authors have also investigated the problem of clustering consumers into groups by their consumption pattern. Knowledge about the characteristics of such clusters can be used to develop novel tariff schemes, improve network management, or to perform load forecasting. An early example of an approach that clusters consumption data from a large number of households is provided by Chicco et al. [20]. The authors cluster electricity consumption data of 471 non-residential customers. Evaluating the clusters along with the current tariffs of each of the customers, they detect examples of inefficient billing practices (e.g., in case there is a poor correlation between discriminatory factors and actual load patterns) [20]. Many approaches related to this use so-called self-organizing maps (SOMs) to cluster a large number of different households based on their electricity consumption [21, 22, 23]. For instance, Figueiredo et al. use this type of unsupervised learning to identify clusters of households with similar consumption behavior [21]. The authors further create rules that form a decision tree to automatically assign new households to one of the clusters by following that tree. Their results are based on electricity consumption traces from 165 households in Portugal collected at a 15-minute granularity. Other researchers also include different type of information to the analysis of plain electricity consumption data [22, 23]. Sanchez et al. [22], for instance, compute specific features out of the data and feed these features to a self-organizing map (SOM) along with additional information obtained through questionnaires. Instead of using electricity consumption to train the SOM, Räsänen et al. rely on properties of the dwelling [23]. This way, the authors make households in the same cluster comparable with respect to their dwelling characteristics. Ultimately, identifying the households with a high (annual) electricity consumption in each cluster, Räsänen et al.’s method allows to create targeted energy saving recommendations.

In a previous study, we also relied on SOMs to cluster groups of households based on their electricity consump-

tion [12]. In contrast to the SOM-based approaches presented above, we also evaluated the correlation of household properties with electricity consumption data. We performed this study to identify properties that have a high correlation and are considered particularly useful by domain experts – which we learned by conducting interviews with energy consultants from four different energy providers in Switzerland. In this paper we now investigate these properties in detail using a supervised learning method.

McLoughlin et al. analyze the correlation between the electricity consumption of a household and its demographic, socio-economic, and dwelling characteristics [24]. The data set used in this study is significantly larger than others previously considered and – although not explicitly stated – is most likely to be the same data set that we use for our investigation. Applying a multiple linear regression model on the data the authors show that certain properties (e.g., number of bedrooms, age of the head of household) have a large influence on the electricity consumption. The authors further show a relationship between maximum demand and the penetration of certain household appliances. Kolter et al. utilize such correlations and propose a regression method that leverages knowledge of dwelling properties such as size, insulation, or location, to estimate the most likely energy consumption level of a household based on its characteristics [25]. This estimation can then be compared against a household’s actual consumption – as well as against that of similar households.

Our work – in contrast to all other approaches described in this section – estimates household properties by analyzing the household’s electricity consumption data. To this end, we rely on a supervised learning method, using household properties along with the consumption data to train our system.

### 3. SYSTEM DESIGN

The upper part of figure 1 shows the two main components of our classification system: *feature extraction* and *classification*. The feature extraction component takes as input the electricity consumption trace of a household and computes over it a set of *features*. The classification component takes these features as input and uses them to classify the household according to previously specified properties. As depicted in the lower part of figure 1, the design of CLASS comprises three steps, which are explained in the following sections in detail. In section 3.1, we describe the set of features that is computed during the feature extraction. Sections 3.2 and 3.3 address the design of the classification component. First, we define *class labels* from which a classifier takes a possible estimate for the household properties. Second, we choose a set of classification algorithms that perform the classification and motivate our choice.

#### 3.1 Feature Extraction

CLASS takes as input an electricity consumption trace of a private household. In its current implementation, CLASS assumes this trace refers to one week (including weekend) and that one sample every 30 minutes is available. Nonetheless, CLASS can be easily adapted to operate with data of different granularity. One week of data at a 30-minute granularity implies that an input trace contains 336 data samples. In [12], we have already defined a comprehensive set of features that can be computed on electricity consumption

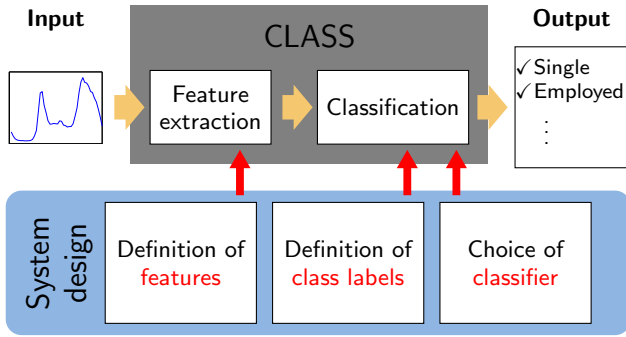


Figure 1: Overview of the design of our classification system. The upper part of the picture shows CLASS’s main components: *feature extraction* and *classification*. The bottom part shows the design steps necessary to specify these components.

data. For its definition, we have considered both existing work in electricity consumption data analysis [20, 21, 22] as well as own considerations relative to the problem at hand. In this study, we use the same set of features as input of CLASS’s classification algorithms.

The 22 features can be divided in four groups: consumption figures, ratios, temporal properties, and statistical properties. Each group contains 9, 5, 4, and 4 features, respectively. Consumption figures are aggregates of the electricity consumption data at different periods of the day (e.g., in the morning from 6 a.m. to 10 a.m. or during weekends). These features allow to compare households with respect to their average electricity consumption during these periods. Ratios are features calculated as the ratio of two different consumption figures (e.g., average consumption in the evening and during lunch time). These ratios allow to capture relevant patterns like whether cooking typically takes place over lunch-time, in the evening, or both. Temporal features capture the time instant of the first occurrence of an event, such as the daily electricity consumption exceeding a specific threshold. Finally, statistical properties include features such as the variance or the total number of peaks of the consumption trace over a day. For more details about the feature set used in this study, we refer the reader to [12].

### 3.2 Class Labels

As we rely on supervised machine learning methods, the classification requires a set of class labels, from which the classifier takes a possible estimate for the label of an unclassified household. The left column of table 1 shows the list of twelve properties considered in this study. The right column of table 1 shows the class labels we have accordingly defined for each property. The class labels represent the possible values that an estimate of a household property can take. In other words, given the consumption trace of a specific household, CLASS will return as output an estimate of the value of each of the properties listed in table 1. In this way, we turn the problem of estimating values of household properties into a classification problem.

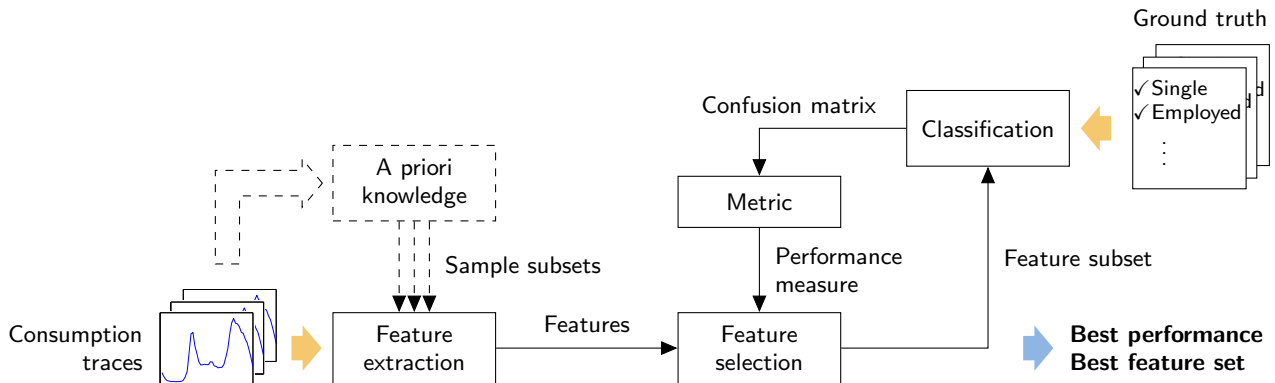
We have compiled the set of properties listed in table 1 using the results presented in our previous work [12]. For example, we have carried out in-depth interviews with energy consultants of four different Swiss energy providers. The

Table 1: List of household properties utilized by CLASS, and their corresponding class labels.

Household property	Classes and their labels
Number of appliances and entertainment devices ( <b>#devices</b> )	Low ( <b>#devices</b> ≤ 8) Medium (8 < <b>#devices</b> ≤ 11) High (11 < <b>#devices</b> )
Number of bedrooms ( <b>#bedrooms</b> )	Very low ( <b>#bedrooms</b> ≤ 2) Low ( <b>#bedrooms</b> = 3) High ( <b>#bedrooms</b> = 4) Very high (4 < <b>#bedrooms</b> )
Type of cooking facility ( <b>cooking</b> )	Electrical Not electrical
Employment of chief income earner ( <b>employment</b> )	Employed Not employed
Family ( <b>family</b> )	Family ( <b>#adults</b> > 1 and <b>#children</b> > 0) No family
Floor area ( <b>floor_area</b> )	Small ( <b>floor_area</b> ≤ 100 m <sup>2</sup> ) Medium (100 m <sup>2</sup> < <b>floor_area</b> and <b>floor_area</b> ≤ 200 m <sup>2</sup> ) Big (200 m <sup>2</sup> < <b>floor_area</b> )
Children ( <b>#children</b> )	Children ( <b>#children</b> ≥ 1) No children ( <b>#children</b> = 0)
Age of building ( <b>age_house</b> )	Old (30 < <b>age_house</b> ) New ( <b>age_house</b> ≤ 30)
Number of residents ( <b>#residents</b> )	Few ( <b>#residents</b> ≤ 2) Many (3 ≤ <b>#residents</b> )
Single ( <b>single</b> )	Single ( <b>#adults</b> = 1 and <b>#children</b> = 0) No single
Retirement status of chief income earner ( <b>retirement</b> )	Retired Not retired
Social class of chief income earner according to NRS social grades ( <b>social_class</b> )	A or B C1 or C2 D or E

goal of the interviews was to derive a list of properties consultants need to know in order to: (1) identify customers potentially interested in energy consulting; (2) prepare for energy consulting sessions. This application-driven approach for identifying interesting properties has for instance outlined that knowing whether a household is occupied by a single person or a family is particularly relevant to consultants. Accordingly, the list in table 1 includes the properties **single** and **family**. For the evaluation of our system, it is also necessary to verify whether the ground truth of the desired household property is available in the CER data set. For instance, the property **family** requires information about the number of adults (**#adults**) as well as the number of children (**#children**) living in a specific household. Similarly, **#residents** describes the number of persons living in the household. The availability of this ground truth data allows us to verify whether and with which accuracy CLASS can classify a household as belonging to the class *Single* / *No single* or *Family* / *No family*. Similar considerations apply to all other properties listed in table 1.

We should note at this point that while for some of the properties there exists a “natural” definition of class labels (e.g., *Single* / *No single*), for other properties – like property **age\_house** – several different sets of class labels can



**Figure 2: Evaluation system used to assess the performance of CLASS. The system allows to evaluate the performance of all considered classifiers in terms of their ability to estimate unknown household properties from electricity consumption data. The system also computes the feature set that allows to obtain the best classification results. Optionally, the system can also integrate a priori knowledge.**

be defined. For these properties, we define two (or more) class labels according to the following criteria. Either we define labels such that the number of households per class is distributed (roughly) equally or we define labels according to qualitative considerations gathered during the aforementioned interviews. Class labels corresponding to the properties `#devices`, `#residents`, and `social_class` have been defined using the first criterion. Class labels for properties `#bedrooms`, `age_house`, `floor_area` using the second.

### 3.3 Classification

As mentioned before, the design of CLASS relies on supervised machine learning methods to classify the households. To this end, CLASS integrates a set of classifiers that “learn” the correlation between a given input feature vector and the corresponding (correct) class label. Hereby, a subset of input data for which the actual class labels are known is used to train the classifier and evaluate its performance. The performance is based on the number of correctly predicted class labels for each of the properties listed in table 1. For each property, CLASS then relies on the classifier that predicted most labels correctly in order to estimate the class label of an unclassified household.

There is a variety of classification algorithms available in literature (see e.g., [26] for an overview) that can be used within CLASS. The choice of classifiers for our system depends on several factors including ease of implementation, computational complexity, and achievable classification accuracy. In order to obtain a comprehensive picture of CLASS’s performance, we have selected four well-known, representative classifiers: the  $k$ -Nearest Neighbor (kNN) classifier [26], the Linear Discriminant Analysis (LDA) classifier, the Mahalanobis classifier [27], and the Support Vector Machine (SVM) [28].

For a detailed description of the four classifiers, the reader is referred to [26, 27, 28]. Here, we outline the specific trade-offs exposed by these classifiers and provide few details about their implementation within CLASS. Besides its simplicity, a main advantage of the kNN classifier is that it does not make any assumption on the distribution of the input data, which also does not need to be linearly separable. On the other side, the kNN classifier has high computational and memory requirements. For the LDA classifier,

we assume a (multivariate) Gaussian distribution of the input data samples. This causes the parameters of the discriminant functions that partition the feature space to be dependent on the mean and covariance of the distributions for each class. The linear functions of the LDA classifier are obtained by assuming a common (pooled) covariance matrix for all classes, constructed by averaging the covariance matrices of each class. The need to assume Gaussianity of the input data is a major drawback of the LDA classifier. On the other side, it has very low requirements in terms of computation and memory usage. The Mahalanobis classifier is conceptually similar to the LDA classifier. One of the main differences is that the former relies on stratified covariance matrices, instead of a pooled covariance matrix. This results in quadratic discriminant functions, which typically allows the Mahalanobis classifier to have better classification performance than the LDA classifier. However, its performance is also more sensitive to the estimation accuracy of the stratified covariance matrices. SVMs are widely used in classification applications [26], which is due to their flexibility and thus, applicability to many classification problems of different natures. A major strength of SVMs is their ability to compute decision boundaries without assuming specific distributions of the input data (like the kNN classifier but unlike the LDA or Mahalanobis classifiers). Further, SVMs are able to cope with data that is not linearly separable, since they support non-linear decision boundaries. A major drawback is the computationally expensive training phase.

## 4. EVALUATION SETUP

After having discussed CLASS’s design in the previous section, we now discuss how we evaluate its classification performance. In particular, we introduce the *metrics* used to quantify this performance and discuss the use of *feature selection algorithms*. Furthermore, we outline how the use of *a priori information* can help improving CLASS’s overall performance. Figure 2 depicts our evaluation setup.

### 4.1 Metrics

The most common metric used to evaluate the performance of a classifier is the *accuracy* [29, 30]. For a two-class classification problem, this metric is computed as a function

of four other quantities: the *true positives* (TP), *true negatives* (TN), *false positives* (FP), and *false negatives* (FN). Given a “target” class A and another class B, TP indicates the number of samples of A that are correctly classified as A. TN denotes the number of samples of B that are correctly classified as B. FP counts the number of samples of B that are incorrectly classified as A and, finally, FN indicates the number of samples of A that are incorrectly classified as B. These four values form the so-called *confusion matrix*, which is schematically depicted in figure 3.

		Predicted		Total
		Class A	Class B	
True	Class A	TP	FN	$S_A = TP + FN$
	Class B	FP	TN	$S_B = FP + TN$

**Figure 3: Generic confusion matrix.**

The accuracy of a classifier is defined as the ratio between the number of correct classifications and the total number of test samples [30]. Using the notation introduced above it can be computed as:

$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (1)$$

In a general classification problem with  $K$  classes ( $K$ -class problem), the confusion matrix consists of  $K^2$  elements. The accuracy is then computed as the ratio between the sum of the  $K$  diagonal elements (i.e., the number of correct classifications) and the sum of all test samples.

We compare the accuracy of CLASS to that of two simple classifiers to which we refer to as the *Random Guess* (RG) and *Biased Random Guess* (BRG) classifiers. In a  $K$ -class problem, the RG classifier assigns the input data sample uniformly at random to one of the  $K$  possible classes. The accuracy of the RG classifier is thus  $\text{ACC}_{\text{rg}} = \frac{1}{K}$ . The BRG classifier biases its random decision so that the proportion of households assigned to each of the available classes is equal to the actual distribution of the input data. If  $S_k$  denotes the number of test samples of the  $k$ -th class of a  $K$ -class problem and  $S$  is the total number of test samples, the accuracy of the BRG classifier is computed as:  $\text{ACC}_{\text{brg}} = \sum_{k=1}^K \left(\frac{S_k}{S}\right)^2$ .

We further evaluate the performance of CLASS in terms of *precision* and *recall* [29]. In a two-class problem, precision is defined as

$$\text{PRC} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

and denotes the probability that a sample classified as class A truly belongs to class A. For instance, a precision of 80% in the classification of the property `single` indicates that if a household is classified as a single-person household, it is truly a single-person household with 80% probability. Recall is defined as

$$\text{RCL} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

and denotes the probability of a sample being classified as class A given that the sample belongs to class A. For instance, a recall of 80% in the classification of the property `single` indicates that 80% of the households that are actually single-person households are also classified as such.

Considering the values of precision and recall separately might be misleading. Indeed, a classifier that assigns all test

samples to a class A – irrespectively of the true membership to class A or B – has a recall of 100%. Most likely, however, the same classifier will also have very low precision. Hence, metrics that combine both, precision and recall, are often used when evaluating the performance of a classifier. A commonly used metric is the  $F_1$  score, which is defined as [29]:

$$F_1 = 2 * \frac{\text{PRC} * \text{RCL}}{\text{PRC} + \text{RCL}}. \quad (4)$$

We should note that since precision, recall and  $F_1$  score assume that there exists a “target” class, they can only be computed for two-class problems. Accuracy can instead be computed for generic  $K$ -class classifiers.

## 4.2 Feature Selection

As shown in figure 2, the feature set described in section 3.1 and computed by CLASS’s feature extraction component, undergoes a *feature selection* step before being used for the actual classification task. We include this step in our evaluation since using the complete feature set as input data might result in suboptimal classification performance [31]. Thus, while training the classifiers we search for the *optimal feature set*, i.e., the subset of features that maximizes classifiers’s performance. This optimal set can be determined by performing an exhaustive search that examines all possible subsets of the feature set [31]. The subset that provides the best overall performance is then selected as the optimal feature set. However, the computational complexity of such an exhaustive search grows exponentially with the number of features [32]. So-called *feature selection algorithms* can thus be used to approximate an exhaustive search, although the optimal set provided by these algorithms might differ from the “true” optimum. CLASS implements two feature selection algorithms: *Sequential Forward Selection* (SFS) [31] and *Parallel Sequential Forward Selection* (P-SFS) [33, Sec. 4.2]. P-SFS typically performs better than – or at least as good as – SFS but requires slightly more computational and memory resources. The rationale behind both these algorithms is to avoid an exhaustive search by iteratively increasing the size of the feature set. Thereby, only the feature that – at each iteration – allows to maximize a given performance metric, is retained. In this study, we use accuracy as the performance metric of choice, although any other of the metrics mentioned above could also be used. For a more detailed description of SFS and P-SFS we refer the reader to [31] and [33, Sec. 4.2], respectively.

## 4.3 A Priori Knowledge

In some settings, partial information about household properties – e.g., their floor area – might be known a priori. Such information can for instance be obtained from public tax record repositories (as done, e.g., by Kolter et al. [25]) or through dedicated customer surveys. We thus evaluate the performance of CLASS also assuming that the values of selected properties are known in advance. Integrating this a priori information in the evaluation of CLASS allows to measure whether knowledge of selected properties might be valuable or not to improve CLASS’s performance. Providers could then accordingly take action to retrieve the information that can be collected at lower cost and allows for the highest performance gains.

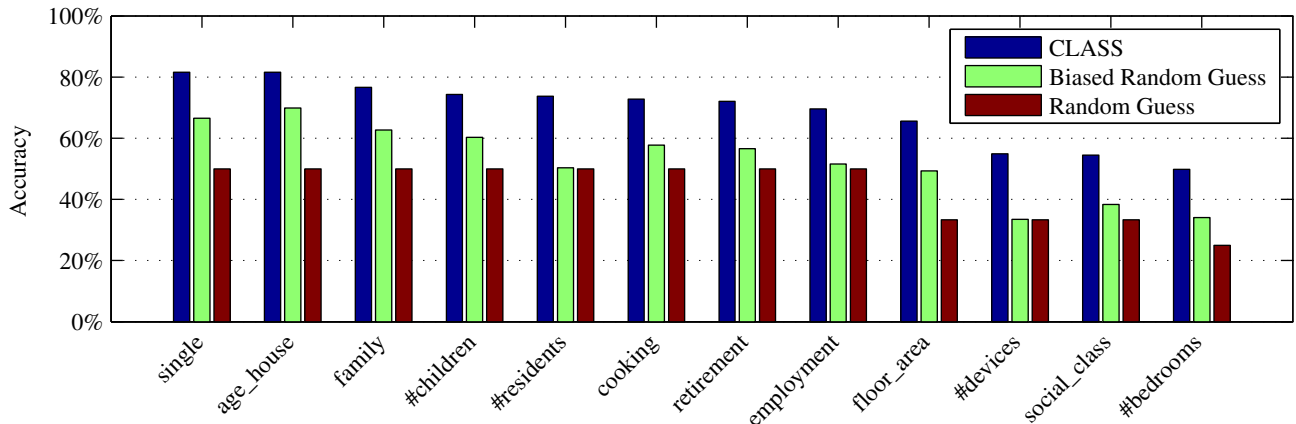


Figure 4: Accuracy of CLASS compared to the random guess and biased random guess classifiers.

#### 4.4 Implementation Details

The evaluation of CLASS relies on 3,488 out of the 4,225 electricity consumption traces of private households included in the CER data set. The 737 neglected data traces are those for which questionnaire answers were not available or participants withdrew from the trial beforehand. From the 1.5 years described by the data we selected consumption traces from week 26, which is the second week of January (Jan 11, 2010 – Jan 17, 2010). We also run our experiments on all the remaining 72 weeks, showing that results obtained for week 26 are representative for the whole data set. We used the data of this week both for training and testing. To avoid data being part of both the training and test set at the same time, we use four-fold cross-validation. On average, the traces of 2,616 households are used for training and the remaining 872 households for testing. Each household is thus used three times for training and once for testing.

The results presented in section 5 are obtained using our own Matlab implementation of CLASS. We evaluate the classification performance for each of the twelve properties listed in table 1. All experiments are repeated independently for each of the four classifiers (kNN, LDA, Mahalanobis, SVM) described in section 3.3. For the implementation of the kNN, LDA, and Mahalanobis classifiers we use Matlab’s Statistics toolbox<sup>2</sup>. For kNN, we choose  $k = 5$  and use the Euclidian distance as the distance metric. We further employ a k-d tree for space partitioning to perform the neighbor search to avoid a linear search. As for the SVM classifier, we rely on the publicly available C-SVM implementation LIBSVM<sup>3</sup>. Setting  $C = 1$ , we follow a common choice to penalize outliers in the training data. For the P-SFS algorithm, we run three branches in parallel.

### 5. CLASSIFICATION RESULTS

In this section, we present the classification performance achieved by CLASS using the CER data set. In particular, we present results in terms of accuracy,  $F_1$  score, precision and recall. Furthermore, we analyze the influence that the availability of a priori information has on the classification

performance. All experiments have been carried out using the setup described in the previous section.

#### 5.1 Accuracy

Figure 4 shows the accuracy of CLASS for all the twelve household properties listed in table 1. For each property, the accuracy obtained by CLASS is compared to that of the Biased Random Guess (BRG) and Random Guess (RG) classifiers described in section 4.1. The accuracy of CLASS is selected as the highest among those obtained by the four classifiers described in section 3.3 (kNN, LDA, Mahalanobis, and SVM). We postpone the discussion of the accuracy obtained by the single classifiers to section 5.3. Please note that – as indicated in table 1 – the property #bedrooms can assume four different values. Thus, the estimation of this property corresponds to a four-class classification problem. Accordingly, the estimation of properties floor\_area, #devices, and social\_class is a three-class problem and that of the eight remaining properties a two-class one. For these “two-class properties” (the first eight from left in figure 4), the accuracy of CLASS ranges from 70% (employment) to 82% (single and age\_house). For the three-class properties, the accuracy varies between 55% (social\_class) and 66% (floor\_area). Finally, the accuracy obtained for the only four-class property (#bedrooms) is 50%.

The accuracy of the Random Guess classifier ranges between 25% and 50% and that of the Biased Random Guess classifier between 34% and 70%. Overall, the performance increase – in terms of accuracy – obtained by CLASS with respect to the RG and BRG classifiers is 20%–32% and 12%–23%, respectively. The better performance of the BRG classifier is due to the fact that it assumes prior class probabilities to be known. CLASS can thus estimate with high accuracy ( $> 80\%$ ) the properties single and age\_house and achieves good performance (accuracy  $\geq 70\%$ ) for the other six two-class properties (family, #children, #residents, cooking, retirement, employment). This implies that electricity providers can estimate with good accuracy the values of selected properties of their customers’ households by analyzing their electricity consumption data. Providers can then in turn use this information to develop and deploy premium services for their customers. For instance, knowledge of the values of the properties age\_house and #devices

<sup>2</sup>www.mathworks.de/products/statistics

<sup>3</sup>www.csie.ntu.edu.tw/~cjlin/libsvm

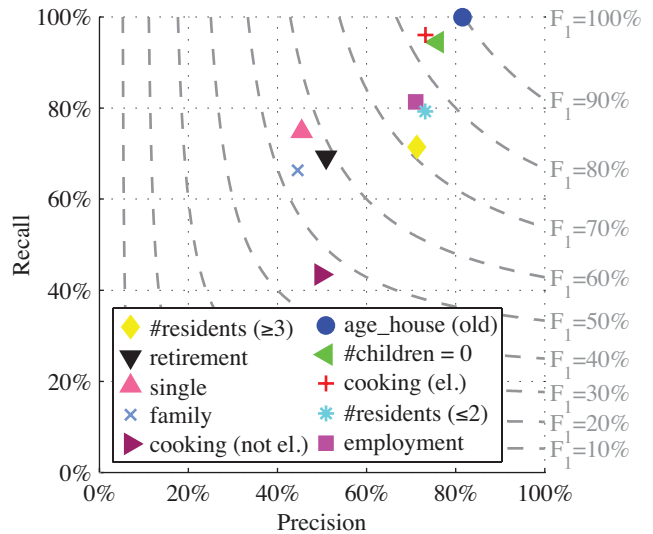
can be useful to prepare consulting sessions and elaborate household-specific energy saving tips.

We recall that the results presented in this section have been computed using one week of electricity consumption data as input (week 26 of the CER data set). To verify that the obtained results are representative for the whole CER data set, we assessed CLASS’s performance using the consumption traces of all of the 72 weeks of the CER trial. For the property `single`, for instance, the resulting standard deviation of the accuracy is 0.7%, 1.2%, 1.3%, and 0.6% for the kNN, LDA, Mahalanobis, and SVM classifier, respectively. Furthermore, the accuracy obtained for this property with input data from week 26 is – with respect to that obtained for all other 72 weeks – the 38th highest for kNN, 45th highest for LDA, 71st for Mahalanobis, and 24th for SVM. Similar results apply to all other household properties considered in this study. Given these considerations, we conclude that there is no loss of generality in presenting results relative to one week of data only (presenting all results is impractical due to space constraints).

## 5.2 Precision and Recall

Figure 5 shows the values of precision and recall of all two-class properties listed in table 1. For the property `#residents` both its class labels *Few* ( $\leq 2$ ) and *Many* ( $\geq 3$ ) are considered, as this information is particularly interesting to energy consultants. Similarly, the precision and recall of both class labels of the property `cooking` (*Electrical* / *Not electrical*) are shown. The further right on the plot a property is located, the higher its precision; the further on the top, the higher its recall. Accordingly, the closer a data point is to the top right corner, the higher is the  $F_1$  score (see equation 4) of the corresponding property. To improve the readability of the plot, we have also plotted 10 different level curves (from  $F_1 = 10\%$  to  $F_1 = 100\%$ ) of the  $F_1$  score. For each property, the figure shows the precision and recall values corresponding to the highest  $F_1$  score among those achieved with the four classifiers used within CLASS.

The graph shows that, for instance, CLASS achieves the best performance ( $\sim 90\%$ ) in terms of  $F_1$  score when classifying households according to property `age_house` and in particular for the class label *Old*. The recall in this case is 99.89% and the precision 81.61%. This implies that if the age of the building to which the household belongs to is actually *Old* (i.e., older than 30 years, as indicated in table 1), then CLASS will classify the household correctly in 99.89% of the cases. On the other side, if CLASS classifies a household as *Old*, then in 81.61% of the cases the household is actually *Old*. Similar considerations apply to other properties with a high  $F_1$  score, like `#children` and `cooking` for the class labels *No Children*, and *Electrical*, respectively. Figure 5 also shows that households with less than three residents (property `#residents`, label *Few*) and those having an employed chief income earner (property `employment`, label *Employed*) are classified with a precision of 71% and 73%, respectively, and with a recall of approximately 80%. The recall value implies that if an electricity provider is interested in selecting, for instance, the specific group of households that have only few (i.e.,  $\leq 2$ , as indicated in table 1) residents, then CLASS would correctly identify 80% of all households of this type. In other words, the provider would be able to address 80% of the customers that live



**Figure 5:  $F_1$  score obtained for all two-class properties. Each marker shows the highest value among those obtained by the considered classifiers.**

in households with few residents. This can be particularly useful to elaborate, e.g., personalized marketing campaigns.

Relatively low  $F_1$  scores are obtained for properties `retirement`, `single`, and `family`. In particular, the precision for the property `single` is just roughly 45%. This means that if CLASS classifies a household as *Single*, then in only 45% of the cases this classification is correct. On the other side, the recall value for the property `single` (label *Single*) is 75%, i.e., 75% of the households occupied by a single person are also classified as *Single*. One of the reasons for this low precision is that only 27% of the households in the CER data set are actually single-person households. As in any classification problem, such “rare samples” are hard to detect. The fact that the accuracy for the property `single` is 80% (as discussed in section 5.1) shows the importance of considering the performance of a classifier also in terms of precision and recall. The classifiers can indeed be tuned so that the one or the other metric is maximized. The results discussed in this section have been obtained by using accuracy as the figure of merit to maximize during feature selection. However, if the goal of the classification is to identify households belonging to a target group (e.g., single-person households), then the classifiers must be tuned to maximize the  $F_1$  score. Evaluating the change in performance obtained when configuring CLASS to maximize the  $F_1$  score is a part of our future work.

## 5.3 Classifier Comparison

Figure 6 shows the performance in terms of accuracy for each of the four classifiers implemented in CLASS and for each of the considered household properties. Figure 7 reports the corresponding  $F_1$  score. The SVM classifier provides the highest accuracy for eight of the twelve household properties and the second highest for the remaining four properties. This strength of the SVM classifier is due to its ability to classify properties even in presence of non-linearly separable input data. It becomes particularly vis-



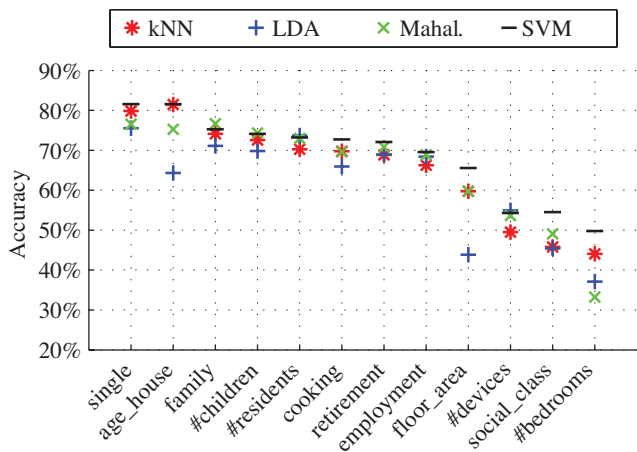


Figure 6: Accuracy of different classifiers for all household properties.

ible for multi-class properties (`floor_area`, `#devices`, `social_class`, and `#bedrooms`) as SVM clearly outperforms the other classifiers. LDA shows the worst performance for seven out of twelve household properties. For the properties `age_house` and `floor_area`, the accuracy provided by the LDA classifier is significantly lower than the accuracy provided by the three other classifiers. LDA’s poor performance is due to the fact that the input data is not normally distributed (but it is assumed to be so by LDA). The kNN and Mahalanobis classifiers provide performances close to those of SVM for most properties but the multi-class ones.

The performances of the classifiers in terms of  $F_1$  score are less homogeneous. Figure 7 shows that the LDA classifier provides the highest  $F_1$  score for five out of ten properties. However, it also performs worst for four properties. Similar considerations apply to the performance of the other classifiers. When classifying the property `single`, for example, the  $F_1$  score of the kNN and SVM classifiers is significantly lower than that of the LDA and Mahalanobis classifiers. We believe that the differences in classification performance with respect to the  $F_1$  score can – at least partially – be attributed to the different number of households available for each class and property. In particular, the LDA classifier outperforms the other classifiers in terms of  $F_1$  score when the classification requires identifying households that are outnumbered by the households of other classes. Hence, LDA is particularly good at detecting rare samples in the CER data set.

The results show that the choice of a particular classifier depends on the specific target application. In general, SVM stands out as the classifier that, for the given data set, allows to achieve the best performance in terms of accuracy. However, other classifiers might allow for slightly better performance for selected properties (e.g., Mahalanobis for `family`). If the application involves identifying households of a specific group, the classifier should be selected based on application-specific requirements (e.g., low false positive rates) as well as on the estimated number of households per class.

## 5.4 A Priori Knowledge

As already mentioned in section 4.3, the use of a priori knowledge can help increasing CLASS’s classification per-

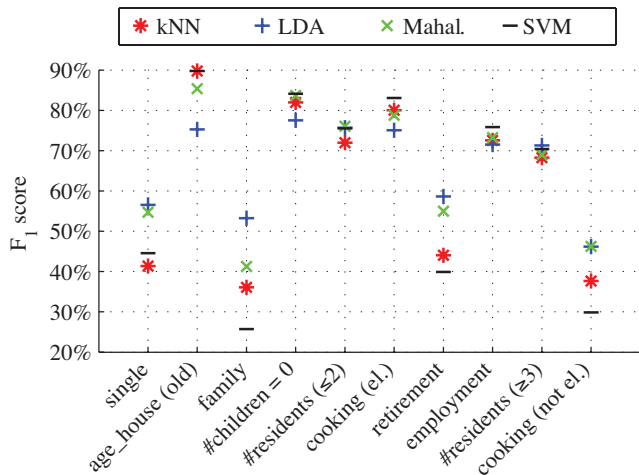


Figure 7:  $F_1$  score of different classifiers for all household properties.

formance. Due to space constraints, we present here results obtained for two representative scenarios only. First, we investigate how the accuracy of the property `single` changes when information about other properties is available. In other words, we analyze which information should preferably be gathered (e.g., through customer surveys) in order to improve the accuracy of the property `single`. Second, we analyze how the accuracy obtained by CLASS for different properties changes when knowledge of one specific *a priori property* is known. In particular, the a priori property used below indicates the age of one of the residents and is indicated as `age_person`.<sup>4</sup>

The black segments in figure 8 show the baseline accuracies achieved when classifying property `single` when no a priori information is available. In contrast to section 5.1, we restrict the set of households to the ones where a priori information is provided by the questionnaires. This is noteworthy as property `floor_area` is only provided by a subset of households. Classifying property `single` on this subset results in a significantly higher accuracy (86%) compared to the subsets denoted by the other a priori properties (82%), which contain almost all households in the data set.

We then run CLASS to determine how the accuracy changes when each of the a priori properties listed on the x-axis is assumed to be known. To this end, we first select households belonging to a known class and then run the classification on this reduced data set. For instance, one of the a priori properties we consider for this experiment is `floor_area`. We select households with a floor area bigger than  $180\text{ m}^2$  and run CLASS over the data corresponding to these households. The cross labeled as  $> 180\text{ m}^2$  in figure 8 shows the accuracy obtained in this experiment. Accordingly, the cross labeled as  $< 180\text{ m}^2$  shows the accuracy obtained when households with a floor area smaller than  $180\text{ m}^2$  are considered as input data set. The stern marker in between the two crosses indi-

<sup>4</sup>The CER data set contains information about the age of the person who filled in the questionnaires for each household. From this information, we derive whether a person younger than 35, older than 65, or in between 35 and 65 years old is a resident of the household.

cates the mean accuracy weighted by the number of samples of the two classes ( $> 180 m^2$  and  $< 180 m^2$ ). Hence, this value represents the overall accuracy that is achieved when the property `floor_area` is known a priori. The same procedure is applied to obtain the accuracy when other properties and specific class labels are given. Other properties include `age_person` (class labels as defined above), `#bedrooms` (class labels 1–3 and  $> 3$ ), `ownership` (which describes whether or not the residents of a household own the house – class labels *Own* and *Rent*), `social_class` (as defined in table 1), and `house_type`, which indicates whether the residents’s house is *Terraced*, *Detached*, *Semi-Detached*, or a *Bungalow*.

Figure 8 shows that when the property `floor_area` is known a priori the accuracy in the classification of a household as *Single* or *Not single* increases by 5.1% to 91% for households with floor area larger than  $180 m^2$ . However, it decreases by 1.9% to 84% for households with floor area smaller than  $180 m^2$ . The average accuracy increase with respect to the case in which no a priori knowledge is given is 0.7%. A similar effect can be observed when knowing property `#bedrooms` a priori. Restricting the set of households to the ones with more than three bedrooms increases accuracy when classifying property `single` by 7.7% to 89%. On the other side, the accuracy when classifying households with one to three bedrooms is 75.9% and thus 5.4% lower than the baseline. Overall, knowing the number of bedrooms a priori increases accuracy by 0.7%. When `age_person`, `ownership`, `social_class`, or `house_type` are used as a priori properties, the average accuracy changes also vary between  $-0.4\%$  and  $1.6\%$ . Interestingly, knowing whether a house is owned or rented (indicated by the property `ownership`) decreases accuracy in both cases. This effect is possible since the distribution of the data changes when the set of households is separated into multiple subsets.

The results presented in figure 8 allow to draw the conclusion that – considering the accuracy for single-person households – an accuracy increase can also be obtained if it is known whether: at least one person younger than 65 lives in the household (property `age_person`); the household has more than three bedrooms (property `#bedrooms`); the house has more than  $180 m^2$ ; the social class of the chief income earner of the household is *A*, *B*, *C1*, or *C2* (property `social_class`); the household is a (semi-)detached house (property `house_type`). A more detailed analysis considering other properties than `single` is left to future work.

Figure 9 shows the accuracy obtained when classifying the properties `#bedrooms`, `#devices`, `employment`, `floor_area`, `#residents`, `single`, and `social_class` in five different experiments labeled as *None*, *Avg*,  $< 35$ ,  $35 - 65$ , and  $> 65$ . Experiment  $> 65$  refers to the case in which classification is run for the set of households that have at least one resident older than 65. In other words, the experiment is run assuming the `age_person` property (class  $> 65$ ) is known a priori. Accordingly, experiments  $< 35$  and  $35 - 65$  are run using as input data only electricity consumption traces of households having at least one resident younger than 35 or in between 35 and 65, respectively. The results labeled as *Avg* represent the weighted accuracy of the accuracy of experiments  $< 35$ ,  $35 - 65$ , and  $> 65$ . The weight for each experiment is computed as the ratio of number of households included in each experiment and the total number of households in the CER data set for which the `age_person` property is given. Thus, *Avg* indicates the accuracy obtained by using `age_person`

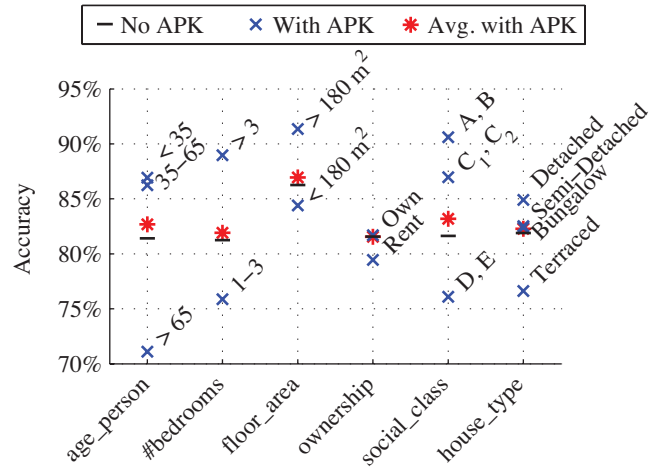


Figure 8: Accuracy obtained by CLASS when classifying the property `single` using different types of a priori knowledge (APK).

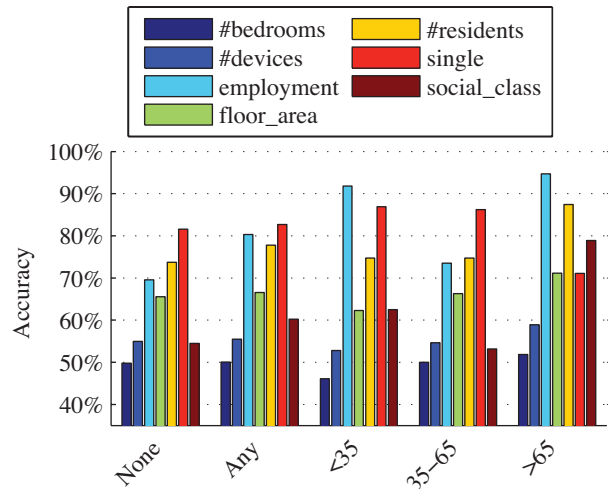


Figure 9: Accuracy obtained by CLASS when the value of the property `age_person` is known a priori.

as a priori property. For the sake of comparison, *All* shows the accuracy obtained by CLASS when all households are included into classification (i.e., no a priori information is available).

The consequences of using `age_person` as a priori property property depend both on the type of a priori knowledge available as well as on the property to be classified. For instance, knowing whether a person in between 35 and 65 lives in the households (experiment  $35 - 65$ ), has almost no effect of the accuracy of the properties shown in figure 9. Instead, significant accuracy gains can be achieved when the a priori information tells whether at least one resident is younger than 35 or older than 65. In the second case, the accuracy of properties `employment`, `floor_area`, `#residents`, and `social_class` increases by 25.7%, 5.5%, 14.0%, and 24.6%, respectively. At the same time, the accuracy of the property `single` decreases by 10.3%. For the case in which it is known that at least one person younger than 35

lives in the household, the accuracy of properties `employment`, `single`, `#residents`, and `social_class` increases by 22.8%, 5.5%, 1.4%, and 8.3%, respectively. However, the accuracy of the property `floor_area` decreases by 3.3%. As for the accuracy of the remaining properties (`#bedrooms`, `#devices`), only marginal changes are obtained. The weighted accuracy (illustrated in figure 9 as *Avg*) can be improved by 11.3%, 4.4%, and 6.0% for properties `employment`, `#residents`, and `social_class` to a total of 80.3%, 77.8%, and 60.2%, respectively.

Considering these results, acquiring information about the age of at least one of the residents of a household is definitely valuable to learn other properties of the same household. Performing classification using all households traces (*Avg*) allows to improve the accuracy of properties `employment`, `#residents`, and `social_class`. For these properties the performance gain can be further strengthened by selecting only households with at least one resident older than 65. However, this additional gain is only achievable for the selected subset of households.

## 6. LIMITATIONS AND OUTLOOK

The paper shows that classification of households based on their electricity consumption is feasible for the household properties we investigated. Though, we expect that there is still room to optimize CLASS and improve classification performance. First, extending the feature set to provide information from more sources or more detailed information to the classifiers can further improve classification performance. For instance, we plan to integrate historical values of the outdoor temperature in Ireland as well as vacation information. We also plan to exploit more statistical properties on the load curves and to compute features for individual days rather than using a weekly average. With respect to feature selection, we will implement the sequential floating feature selection method (SFFS), which is considered as one of the best performing feature selection methods [34]. To increase the performance of CLASS for use cases in which high precision and recall scores are more relevant, we will change the figure of merit during feature selection to the  $F_1$  score. Finally, we will investigate the boosting effect by combining classifiers, where – in the ideal case – some classifiers can compensate for the weaknesses of others.

Besides optimization of precision and accuracy, it is a part of future work to provide more detailed information with the classification results. Until now, for each household property, CLASS estimates the class a household (most likely) belongs to. First, adding a confidence value to each estimation allows energy providers to select households that most likely match a given class. Second, we plan to integrate regression models to estimate continuous variables (e.g., for properties `#residents`, `#devices`, `#bedrooms`) in addition to our class-based analysis.

Future work also contains an evaluation on the requirements for energy providers with respect to the measurement infrastructure. The results of this study are obtained using a single week of 30-minute electricity consumption data. Using data from multiple weeks, both in summer and winter, potentially increases classification performance. However, it also requires more data being measured by the energy providers. Similarly, we will study to what extent classification performance suffers from a decrease in granularity, for instance using hourly, daily, or monthly measurements or

even diluted data. This, in turn, simplifies data collection for energy providers as it reduces data storage and communication overhead. Finally, we investigate the contribution of individual features to classification performance, in order to provide – for each household property – a list of key features that is required (or sufficient) to classify the property.

## 7. CONCLUSIONS

This paper presents the design, development and evaluation of CLASS, a system that classifies private households according to pre-specified properties. Examples of such properties include the floor area of a household or the number of its occupants. CLASS analyzes the electricity consumption data of the household and estimates the values of its properties. We evaluate the performance of CLASS on data from 3,488 households and for twelve different properties. Our results outline that CLASS can classify with good accuracy ( $> 70\%$ ) eight out of these twelve properties. The performance in terms of precision and recall as well as for different types of classifiers is also analyzed. Furthermore, we evaluate the impact that the availability of a priori information has on CLASS’s classification performance. The quantitative results presented in this paper refer to the performance obtained by CLASS on the CER data set and are not necessarily representative also for other data sets. However, our results show that automatic classification of households based on electricity consumption data is feasible.

## 8. ACKNOWLEDGEMENTS

This work has been partially supported by the Hans L. Merkle Foundation and by the LOEWE research initiative of the State of Hesse, Germany, within the Priority Program Cocoon.

## References

- [1] J. Vasconcelos, “Survey of regulatory and technological developments concerning smart metering in the European Union electricity market,” *EUI RSCAS Policy Papers*, no. 1, 2008.
- [2] 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.
- [3] Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC.
- [4] F. Mattern, T. Staake, and M. Weiss, “ICT for green – How computers can help us to conserve energy,” in *1st International Conference on Energy-Efficient Computing and Networking (e-Energy)*, pp. 1–10, ACM, 2010.
- [5] C. Fischer, “Feedback on household electricity consumption: a tool for saving energy?,” *Energy Efficiency*, vol. 1, no. 1, pp. 79–104, 2008.
- [6] S. Darby, “The effectiveness of feedback on energy consumption. A review for DEFRA of the literature on metering, billing and direct displays,” 2006.

- [7] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," in *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 239–240, 2011.
- [8] H. Kim, M. Marwah, M. Arlitt, G. Lyon, and J. Han, "Unsupervised disaggregation of low frequency power measurements," in *SIAM International Conference on Data Mining (SDM)*, pp. 747–758, 2010.
- [9] G. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.
- [10] F. Nogales, J. Contreras, A. Conejo, and R. Espínola, "Forecasting next-day electricity prices by time series models," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 342–348, 2002.
- [11] D. De Silva, X. Yu, D. Alahakoon, and G. Holmes, "A data mining framework for electricity consumption analysis from meter data," *IEEE Transactions on Industrial Informatics*, vol. 7, pp. 399–407, 2011.
- [12] 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)*, pp. 169–176, 2012.
- [13] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.
- [14] Council of the Organization for Economic Co-operation and Development (OECD), "OECD guidelines on the protection of privacy and transborder flows of personal data," 1980.
- [15] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, "Load signature study—Part I: Basic concept, structure, and methodology," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 551–560, 2010.
- [16] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz, "On the accuracy of appliance identification based on distributed load metering data," in *2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT)*, 2012.
- [17] J. M. Abreu, F. P. Câmara, and P. Ferrão, "Using pattern recognition to identify habitual behavior in residential electricity consumption," *Energy and Buildings*, vol. 49, pp. 479–487, 2012.
- [18] S. V. Verdú, M. O. García, C. Senabre, A. G. Marín, and F. J. G. Franco, "Classification, filtering, and identification of electrical customer load patterns through the use of self-organizing maps," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1672–1682, 2006.
- [19] B. Birt, G. Newsham, I. Beausoleil-Morrison, M. Armstrong, N. Saldanha, and I. Rowlands, "Disaggregating categories of electrical energy end-use from whole-house hourly data," *Energy and Buildings*, vol. 50, pp. 93–102, 2012.
- [20] G. Chicco, R. Napoli, P. Postolache, M. Scutariu, and C. Toader, "Customer characterization options for improving the tariff offer," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 381–387, 2003.
- [21] V. Figueiredo, F. Rodrigues, Z. Vale, and J. Gouveia, "An electric energy consumer characterization framework based on data mining techniques," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 596–602, 2005.
- [22] I. Sánchez, I. Espinós, L. Moreno Sarrion, A. López, and I. Burgos, "Clients segmentation according to their domestic energy consumption by the use of self-organizing maps," in *6th International Conference on the European Energy Market (EEM)*, pp. 1–6, IEEE, 2009.
- [23] T. Räsänen, J. Ruuskanen, and M. Kolehmainen, "Reducing energy consumption by using self-organizing maps to create more personalized electricity use information," *Applied Energy*, vol. 85, no. 9, pp. 830–840, 2008.
- [24] 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, 2012.
- [25] J. Kolter and J. Ferreira, "A large-scale study on predicting and contextualizing building energy usage," in *25th Conference on Artificial Intelligence (AAAI)*, 2011.
- [26] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [27] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, 1984.
- [28] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [29] I. Witten, E. Frank, and M. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2011.
- [30] E. Alpaydin, *Introduction to Machine Learning*. MIT press, 2004.
- [31] A. Whitney, "A direct method of nonparametric measurement selection," *IEEE Transactions on Computers*, vol. 100, no. 9, pp. 1100–1103, 1971.
- [32] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 153–158, 1997.
- [33] R. Fandos, *ADAC System Design and its Application to Mine Hunting Using SAS Imagery*. PhD thesis, TU Darmstadt, 2012.
- [34] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognition Letters*, vol. 15, no. 11, pp. 1119–1125, 1994.