

# Can Smart Plugs Predict Electric Power Consumption? A Case Study

Andreas Reinhardt  
School of Computer Science and  
Engineering  
The University of New South Wales  
Sydney, Australia  
andreasr@cse.unsw.edu.au

Delphine Christin  
Institute of Computer Science 4  
University of Bonn  
and Fraunhofer FKIE  
Bonn, Germany  
christin@cs.uni-bonn.de

Salil S. Kanhere  
School of Computer Science and  
Engineering  
The University of New South Wales  
Sydney, Australia  
salilk@cse.unsw.edu.au

## ABSTRACT

The Internet of Things will encompass a rich variety of sensing systems including mobile phones, embedded sensor and actuator systems, and even smart electricity meters. Through their collaborative operation, these billions of embedded devices will realize the vision of smart homes, smart cities, and beyond. The resulting vast amount of sensor data offers unprecedented benefits for society. Smart electricity meters can, e.g., already help utilities monitor the stability of the power grid by periodically reporting each connected household's energy consumption. But there is more, especially when data is available at higher spatio-temporal resolutions, e.g., when distributed smart power meters (sometimes referred to as *smart plugs*) are deployed. In this paper, we assess to what extent these embedded sensing systems can analyze power consumption data in order to make projections of an appliance's future energy demand. In a first step, our system identifies and isolates unique characteristic signatures from collected power consumption traces. Subsequently, time series pattern matching is applied to detect these signatures in real-time data. Based on their occurrence, the appliance's future power demand is predicted. We evaluate our approach with more than 2,500 appliance activity segments collected from 15 different appliance types, and show that accurate forecasts can be made in many cases.

## Keywords

Distributed appliance metering, power consumption prediction, smart buildings, smart grids

## 1. INTRODUCTION

High electricity generation dynamics due to the immediate dependency of renewable sources on environmental parameters are increasingly gaining impact on power grids. One of the core problems is the absence of energy storage components in today's grids. As a result, the surplus energy provided during periods of intense sunshine or high winds can-

not be buffered. Instead, excess energy must be consumed at the time it is being generated—or renewable sources be shut down to not generate energy at all [22]. Similarly, to cover the opposite case of unexpected low yields from renewable sources, spinning reserve plants are constantly kept in operation, although their efficiency is mediocre. We argue that projections of the future power consumption of individual dwellings can be used to mitigate this situation. Reliable forecasts would gain operators additional time to proactively prepare supportive means to maintain the grid's stability. In fact, if the achievable prediction horizon is sufficiently long, the spinning reserve can be downsized because sufficient time will be available to start up supportive plants on demand.

The consistently growing number of sensing devices for home automation may come to the rescue. With an envisioned 50 billion networked devices by the year 2020 [12, 13], a wide distribution of both smart electricity meters as well as intelligent power outlets (*smart plugs*) can be expected [7, 19]. By leveraging the data collection and processing capabilities of these devices, households and businesses can make predictions of their future energy consumption at much finer resolution than the currently used standard load profiles [11]. Sharing these forecasts with the utility companies enables them to draw a fine-grained portrait of the expected power grid status and activate reserve power plants only when necessary. Power consumption predictions are thus vital for the sustainable operation of the future smart grid [15].

In this paper, we present a case study on the use of smart plugs, i.e., distributed data collection and processing devices, to make forecasts of an appliance's power consumption. More precisely, we show how previously observed consumption patterns can be leveraged to predict a device's future demand. To this end, we identify characteristic patterns (*signatures*) in historical data first, which we annotate by the power consumption for the remaining duration of the appliance's activity. Once a database of signatures has been established, our system is ready to make consumption data forecasts. It applies time series pattern matching to detect the presence of the extracted signatures in incoming streams of consumption data. When a signature match has been found, the signature's annotations are used in order to predict the appliance's future power demand and its remaining activity time. In sum, we make the following contributions in this paper:

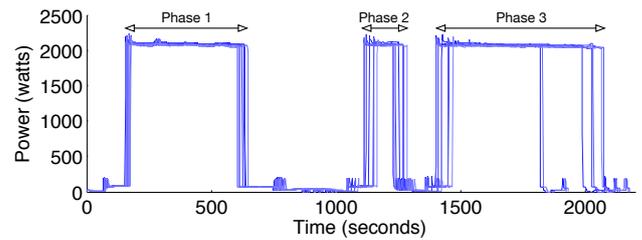
- We present the design of our consumption forecasting system. The system autonomously extracts characteristic signatures from power consumption data and stores all signatures that allow for unambiguous power consumption predictions in a repository.
- We show how time series pattern matching is applied to identify the previously extracted recurring patterns in real-time sensor data streams in order to make consumption predictions.
- Using consumption data collected from real-world devices, we evaluate the achievable prediction accuracy, identify appliance types whose future power consumption can be particularly well determined, and assess the length of the achievable prediction horizon. An evaluation of our system’s operability on embedded systems concludes our evaluation.

This paper is structured as follows. We summarize existing work on predicting power consumptions in Sec. 2, followed by the introduction of our concept and design considerations in Sec. 3. We detail its features and its mode of operation in Sec. 4 and analyze the accuracy of the proposed system in Sec. 5. Finally, we conclude our paper in Sec. 6.

## 2. RELATED WORK

Predicting electric energy consumption, or *load forecasting* for short, has received a lot of attention in the last decade. While early work has mostly targeted to predict the future aggregate power demand observed at substations [1, 24], the increasing deployment of smart meters has allowed more recent approaches to increase the spatial resolution of the forecasts. Several university research groups have, e.g., collected consumption data for their campuses and applied different regression and machine learning techniques in order to extract trends from the data and forecast future usage. Proposed techniques include neural networks [25], support vector machines [9], and decision trees [26]. Furthermore, the combination of historical campus power consumption data with external data sources (e.g., weather information) to improve the results has been presented in [2]. The approaches achieve good accuracy results when a large number of buildings is monitored and time scales of 24 hours or more are applied, because they benefit from the averaging effect of the law of large numbers. Means to make consumption predictions for individual households have been presented in [16], thus increasing the spatial resolution of the forecast even further. However, the approach is based on the manual entry of data from electricity and gas bills and thus only operates on monthly time scales; well in excess of the short-term forecasts targeted in this work.

From the domain of pervasive computing, the prediction of future states has been analyzed for numerous sensing modalities. Some of the resulting approaches have also been transferred to the domain of smart grids in order to predict the electric power consumption of households. To this end, hidden Markov models have mostly been employed to determine the most likely power demand based on a building’s previously encountered power consumption patterns [4, 8]. According to the results presented in these publications, the efficacy of load forecasting without considering additional



**Figure 1: Superimposed power consumption traces of seven operation cycles of the same dishwasher appliance (data taken from [21]).**

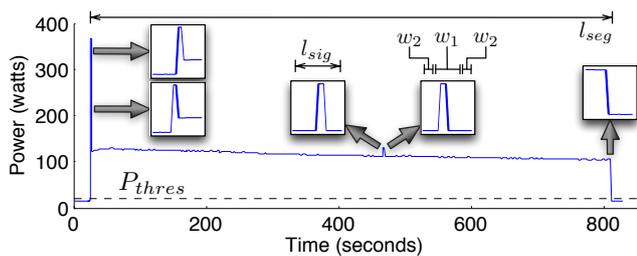
environmental parameters (like building occupancy) is however limited. This observation is confirmed by results on the predictability of individual households that take external parameters into account in order to compensate for the unpredictable user behavior [6]. However, this approach has again been designed to predict a household’s accumulated demand over 24 hours, and does not cater to short-term predictions of load variations on time scales of seconds to minutes. To the best of our knowledge, forecasts for individual households at temporal resolutions of seconds or minutes have not been investigated to date, although the ubiquitous presence of smart plugs and smart meters will enable future buildings to provide data at exactly this resolution.

Infrastructure for the adaptation of supply and demand has been installed on the Danish island of Bornholm within the scope of the European Union’s EcoGrid project [14]. This system however only features loads that monitor the mains frequency and adapt to the current generation situation autonomously. Consumption predictions, such as expected load drops, which might lead to different actuation decisions, are not currently being made by the system.

Our paper shares most similarities with the work presented by Truong et al. in [23], which is also based on fine-grained power measurements. However, while our work targets to forecast the behavior until an appliance is deactivated, the cited publication considers the opposite case, i.e., predicting an appliance’s next activation. Household consumption recordings of several weeks were used in [23] to extrapolate appliance activities and their interdependencies. As the work exclusively relies on historical data, however, the solution is unable to regard the impact of the user’s current behavior on device actuations, and a non-negligible number of false positives were reported in a conducted experiment.

## 3. CONCEPT AND SYSTEM OVERVIEW

Many household appliances are actively controlled by the user, i.e., the user’s physical presence is instrumental to determine when these appliances will be turned *on*. As a result, the sole availability of power measurements makes accurate forecasts of their *activation* time very difficult. In fact, to the best of our knowledge, no related work succeeded in making predictions of appliance activations that are accurate to the minute or even finer temporal resolutions without considering additional environmental parameters. One of the main reasons for this limited predictability is that very few appliances consume measurable power during periods of their inactivity.



**Figure 2:** A refrigerator’s activity segment with five extracted signatures and parameter annotations (threshold power  $P_{thres}$  to facilitate trace segmentation, segment length  $l_{seg}$ , signature length  $l_{sig} = w_2 + w_1 + w_2$ ).

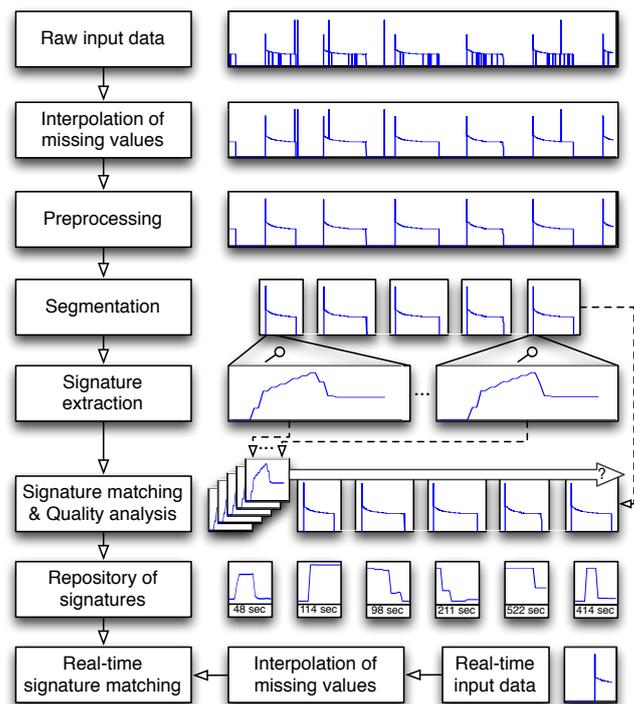
In contrast, the power consumption of almost all devices increases once they have become active. Moreover, recurring consumption patterns can be observed for many appliances. Let us consider the seven superimposed dishwasher operation cycles depicted in Fig. 1, in which three heating phases can clearly be discerned. As the dishwasher’s cleaning program is pre-defined by the appliance manufacturer, future operation cycles can be expected to lead to highly similar time-power relationships. We thus focus on analyzing appliances during their activity intervals, and target to make accurate predictions of their power demand until their deactivation.

### 3.1 System Overview and Terminology

In order to accomplish this task, we have designed a system that extracts and later identifies short characteristic sequences of power consumption data, based upon which it emits power consumption forecasts. We argue that sufficient information to make accurate forecasts is present in an appliance’s past power demand, hence the system does not rely on any additional environmental parameters. In short, it operates as follows.

We first divide the continuous time series of past consumption data into activity *segments*, i.e., time series representations of the appliance’s power consumption during individual phases of its operation. Because each segment has a finite duration, all further extracted characteristics can subsequently relate to the start or end of the segment. In a second step, short characteristic power consumption snippets, so called *signatures*, are extracted from the segments. These signatures are later matched against real-time data in order to make predictions. All signatures are annotated by their distance to the end of the segment and the power consumption until then. For the sake of clarity, we visualize one segment of a refrigerator’s operation in Fig. 2, from which five signatures (shown in the inset boxes) have been extracted.

Before relying on the extracted signatures to make predictions, they must be verified to unambiguously indicate the correct power consumption for the duration of the prediction horizon (e.g., the next 1, 5, or 10 minutes, or the device deactivation—whichever comes earlier). Their extraction is thus followed by a subsequent quality analysis, which eliminates all signatures that would lead to incorrect predictions. Only the remaining unambiguous signatures are eventually



**Figure 3:** Data processing flow of the power consumption prediction system.

stored in a repository, based on which the future consumption behavior of real-time collected data can be forecast. The complete sequence of the steps taken during the system’s initial training phase, including the required preprocessing of the input data, is visualized in Fig. 3 and described in the following subsections.

During the system’s regular operation, incoming power consumption data is matched against the signature repository established in the previous step. This step is shown in the bottom part of the figure. Due to the real-time nature of the input data, however, preprocessing steps that require knowledge of future data points cannot be applied. As a result, only data interpolation is used in order to ensure that the temporal resolution of signatures and input traces match.

## 4. DESIGN AND REALIZATION

In this section, we detail the operation of each of the processing steps shown in Fig. 3. Where applicable, we also introduce the parameter values that have been analytically or empirically defined to provide the best system performance.

### 4.1 Input Data

In order to identify the limits of the achievable accuracy, we conduct our analysis with device-level consumption data. We source our data from the Tracebase project [21], which features appliance-level power consumption traces collected by smart plugs. As opposed to trace collections of circuit-level data (like AMPds [18] or REDD [17]) or even household-wide readings (e.g., BLUED [3]), each trace only contains data of one appliance at a time. Hence, all observed characteristics can immediately and unambiguously be attributed

to an appliance type. Samples are collected at high temporal resolution, such as visible in the traces plotted in Figs. 1 and 2. We are aware of the implications of using device-level data collected by smart plugs and discuss the possible application of our approach at circuit or household level in Sec. 5.3.

For the analysis of the prediction accuracy of our system, we have selected a representative excerpt of the available power consumption traces, using 15 different appliance types commonly found in households. The appliance types span a wide range of consumption levels, from household appliances with a power consumption in excess of 2,000 watts (e.g., washing machines) to lamps that consume less than 20 watts. An overview of the used data is given in Table 1.

## 4.2 Interpolation and Preprocessing

When inspecting the traces, we have observed that the data collection interval of the Tracebase readings is not constant. Although the data set only contains very few traces that have gaps of more than 60 seconds between successive readings, most of the traces have intervals between two and ten seconds between readings. An interpolation of the data has thus proven necessary in order to work with data sets of equal temporal resolution. As we did not want to diminish the quality of the traces that were collected using sampling rates of one or even more readings per second, we have decided to interpolate all traces such that one sample per second is present, i.e., 86,400 samples per day. In order to interpolate the gaps in the data, we have experimentally compared the effects of linear interpolation between the current and the previous data sample against the simple repetition of the last encountered value until a new sample has become available. Our results have suggested the application of the latter solution for two reasons. Firstly, the former method significantly augmented the impact of spurious readings by adding non-existent segments to the traces. Secondly, knowledge about both end points is required for linear interpolation. For streaming live data, where future values are not available, this unnecessarily incurs a time delay and thus reduces the achievable prediction horizon.

In some device traces, power spikes or very short activity periods with durations of only a few seconds have occurred. Even though these readings may be correctly capturing the underlying appliance’s power demand, the prediction of very short sequences is beyond the scope of this paper, in which forecasts in the range of several seconds to minutes are targeted. As a result, we have eliminated all occurrences of consumption periods of less than 20 seconds in order to facilitate the subsequent segmentation of the traces.

## 4.3 Trace Segmentation

After the interpolation and preprocessing steps, 617 traces of 86,400 seconds duration each (equalling more than 1.5 years of input data) remained. A manual segmentation of the traces into individual operational phases would have clearly been very time-consuming. Hence, an automated trace segmentation component follows the preprocessing step. Its task is to extract complete sub-traces that contain single operational cycles from the daily consumption data. For many appliance types, this boils down to extracting individual periods during which the power consumption exceeds the

**Table 1: Statistics of the used input data set.**

Appliance type	# devices	# traces	# segments
Alarm clock	1	5	13
Coffee maker	6	74	141
Cooking stove	1	16	61
Dishwasher	5	74	141
Freezer	1	9	201
Iron	1	3	4
Lamp (halogen)	9	80	99
Microwave oven	7	56	133
Printer	2	12	20
Refrigerator	10	67	1,377
Toaster	4	23	46
Tumble dryer	2	9	16
Washing machine	9	55	74
Water boiler	1	2	42
Water kettle	9	132	421
Total	68	617	2,789

appliance’s standby power demand. Our automated trace segmentation works as follows.

First, a power threshold value ( $P_{thres}$ ) is determined in order to estimate when a transition from standby to active operation occurs, i.e., when a new activity segment begins. Due to the appliances’ different standby power consumptions, we have based this value on the appliance’s power amplitude, i.e., the difference between minimum ( $P_{min}$ ) and maximum ( $P_{max}$ ) encountered power consumption in each input trace according to Eq. (1). The decision in favor of the cube root has been made based on the manual comparison of different means to calculate  $P_{thres}$ , where its usage has resulted in the highest number of correctly separated activity phases.

$$P_{thres} = P_{min} + \sqrt[3]{P_{max} - P_{min}} \quad (1)$$

However, we have also observed that some appliances consume very little power during parts of their activity cycles, e.g., the dishwasher depicted in Fig. 1 between its heating phases. Our segmentation component thus needs to ensure the coherent extraction of complete activity segments, even when they contain power consumption values below  $P_{thres}$ . To this end, we have added a new parameter, the allowed intra-activity idle time  $t_{maxIdle}$ . It bases the allowed idle time during a segment on how long the appliance has been active up to this point. For appliances that have been active for a longer time  $t_{active}$ , a longer idle time is permitted before assuming the end of the activity segment. To ensure its defined operation for appliances with very short operational periods, we have lower-bounded  $t_{maxIdle}$  by an empirically determined parameter  $t_{min}$ . Likewise, the upper limit of  $t_{maxIdle}$  is bounded by  $t_{max}$  in order to reliably detect the end of an activity segment, even when the appliance has been active for a very long time already.  $t_{maxIdle}$  is calculated according to Eq. (2).

$$t_{maxIdle} = \begin{cases} t_{min} & \text{if } t_{active} < t_{min} \\ t_{active} & \text{if } t_{min} \leq t_{active} \leq t_{max} \\ t_{max} & \text{if } t_{active} > t_{max} \end{cases} \quad (2)$$

For most appliances, the parameters have been chosen such that the inter-activity idle time is bounded at  $t_{min} = 300$  seconds and  $t_{max} = 1,800$  seconds. An individual manual review of all input traces and the resulting segments has suggested minor adaptations of these bounds for certain device types in order to ensure correct segmentation. Besides visually confirming that a successful detection of each individual segment was possible, this manual check also ensured that activity cycles comprising idle phases were not split into several parts. Please note that the segmentation of traces is not the core contribution of this paper, and that the user-assisted approach has mainly been chosen because it has allowed us to manually ensure the correct segmentation upon which all further processing steps rely. Hence, it caters to the correctness of the results.

Similar to the removal of all shorts bursts of power consumption in the previous preprocessing step, we have removed all short segments with durations of less than 20 seconds from the data set in order to ensure that long-term predictions can be made. The Tracebase data are stored in chunks of 24 hours duration each, starting and ending at midnight, and very few coherent multi-day traces were available. As a result, we have disregarded possible segments when the first reading in a trace file was already above  $P_{thres}$ . Similarly, we have not extracted segments when an activity started before midnight and did not terminate before the end of the input trace. In summary, we have extracted 2,789 activity segments from the input data, as shown in Table 1.

#### 4.4 Signature Extraction

The next step of our data processing sequence is the extraction of signatures from the segments. Signatures are short snippets of the time series that serve as indicators for an appliance’s subsequent power consumption for a definable period of time (the *prediction horizon*). We consider both the appliance’s prospective deactivation and fixed-length horizons in this paper. To be applicable for making consumption forecasts, signatures need to fulfill two major criteria.

1. Signatures must be unambiguous. In other words, signatures that occur in multiple input traces must be located at identical or very close temporal distance from the end of the segment or, in case of shorter prediction horizons, correctly indicate the power consumption pattern for the entire duration of the chosen prediction horizon.
2. Signatures should be generic, i.e., ideally they should occur in many of the segments. Signatures that have occurred in the training data only once disallow for their generalization, and thus their merit cannot be assessed.

All extracted signatures are comprised of an event window of length  $w_1$ , in which a notable change to the power consumption must occur. Currently, power consumption changes in excess of 10% of an appliance’s amplitude swing (i.e.,  $P_{max} - P_{min}$ ) are considered sufficient to initiate the creation of a new signature. However, a trade-off between the sensitivity of our solution and the number of resulting signatures can be found by adapting the threshold value to the application requirements. We have furthermore lower-bounded the

activity detection threshold to a minimum value of 5 W in order to eliminate errors resulting from the measurement noise observed in the input data. While this makes the approach less suited for small loads, their contribution to a household’s overall load is smaller and thus forecasting their consumption can be considered to be of lower relevance. Alternatively, the use of higher-resolution sensors in smart plugs can be expected to mitigate this problem and thus allow to reduce the value of the lower bound even further.

We surround the event window by two context windows of size  $w_2$ , as depicted in Fig. 2. Throughout all our evaluations, we have chosen  $w_1 = 10$  seconds and  $w_2 = 5$  seconds, hence most segments are 20 seconds in duration. Signatures at the beginning/end of a segment, where no context before/after the event window exists, may however also be shorter and have durations between 15 and 19 seconds only. We annotate all extracted signatures by their distance to the end of their originating activity segment as well as the power consumption until then. Please note that the chosen window sizes are based on our observation that significant power changes usually occur within ten seconds. We consider the investigation of different window sizes and windowing techniques as future work. Based on our approach, a total number of 46,735 candidate signatures were extracted from the input data.

#### 4.5 Signature Quality Analysis

Before the extracted signatures can be used to reliably forecast the future power demand of appliances, confirming their unambiguous prediction capability is of utmost importance. To this end, a signature quality analysis step is applied to all extracted candidate signatures. It checks for the presence of each signature in all activity segments of the same appliance type and removes signatures that fail to fulfill the criterion. As smart plugs are designed to be connected to a single device at a time, it is sufficient for signatures to be unambiguous within their particular appliance class. Nonetheless, we also analyze the effect on the signature repository when signatures from all appliance types are combined in Sec. 5.2. All signatures remaining after the quality analysis step are subsequently added to the signature repository and used to emit predictions for real-time input data of this appliance type.

In order to determine whether signatures are unambiguous, the matching locations of the signature across the segments need to be found. We accomplish this by sliding each candidate signature across all segments of a particular appliance type and calculating the distance between them based on the root mean square deviation  $RMSD$  shown in Eq. (3). In essence, we sum up the squared differences between a signature’s discrete power values  $P_{sig}(t)$  and a segment’s power values  $P_{seg}(t)$  for all time points  $t$  up to the signature’s length ( $l_{sig}$ ), and normalize the result by dividing it by the signature length.

$$RMSD(T) = \sqrt{\frac{\sum_{t=1}^{l_{sig}} (P_{sig}(t) - P_{seg}(T+t))^2}{l_{sig}}} \quad (3)$$

**Table 2: Signature statistics when extracted to make class-wise predictions with different prediction horizons.**

Appliance type	Signature type (60 s horizon)			Signature type (600 s horizon)			Signature type (till deactivation)		
	Unique	Generic	Ambiguous	Unique	Generic	Ambiguous	Unique	Generic	Ambiguous
Alarm clock	3	2	9	2	3	10	3	2	10
Coffee maker	1,762	525	15,282	1,757	479	15,418	1,757	479	15,418
Cooking stove	44	14	41	44	13	44	44	13	44
Dishwasher	508	126	145	511	116	171	508	115	173
Freezer	161	14	57	156	8	83	156	8	83
Iron	88	17	21	88	17	21	88	17	21
Lamp (halogen)	26	14	80	27	11	94	25	10	101
Microwave oven	633	79	81	633	77	85	633	77	85
Printer	173	3	45	173	3	45	173	3	45
Refrigerator	266	107	1,084	254	63	1,370	236	30	1,540
Toaster	26	1	23	26	0	27	26	0	27
Tumble dryer	247	24	29	247	23	32	247	23	32
Washing machine	4,297	156	790	4,297	156	790	4,297	156	790
Water boiler	12	15	67	12	15	67	12	15	67
Water kettle	216	12	223	215	9	229	215	9	229

The value of  $RMSD(T)$  is calculated for all possible signature offsets in the segment ( $T = 0, \dots, l_{seg} - l_{sig}$ ). All resulting offsets  $T$  where the RMS deviation is below a definable threshold level are considered as signature matches. We have chosen a threshold of 2W because it is small enough to avoid ambiguities, but still large enough to compensate for measurement noise introduced by the physical sensors. In order to determine each signature’s unambiguous nature, we analyze for each matching offset  $T$  whether the power consumption prediction to the end of the prediction horizon is sufficiently close (i.e., within 10% of the actual consumption) to the data annotated in the signature. If the appliance deactivation time is being used as the prediction horizon, we have allowed for a temporal difference between observed and predicted appliance deactivation of up to at most 30 seconds.

The resulting numbers of signatures when performing the signature evaluation step on the given input data are shown in Table 2. We have tabulated data for three different prediction horizons of 1 minute, 10 minutes, and the end of the activity segment (i.e., appliance deactivation). Please note that the signature quality assessment has been made on a class-wise basis, i.e., signatures have only been matched against segments of the same appliance type. Values are interpreted as follows.

- *Unique* signatures occur exactly once across all segments. As it is hence unclear whether their predictions can be generalized, these signatures are commonly not added to the repository.
- *Generic* signatures have occurred multiple times in the training data, and at least 90% of their occurrences predict the underlying device’s power consumption until the prediction horizon within the tolerable bounds. They are hence ideally suited to make power consumption predictions in real-time data, and only these signatures are eventually carried over to the final signature repository.
- Signatures in the *ambiguous* column have occurred multiple times, but led to contradictory forecasts within the specified prediction horizon, hence they are not considered for addition to the repository either.

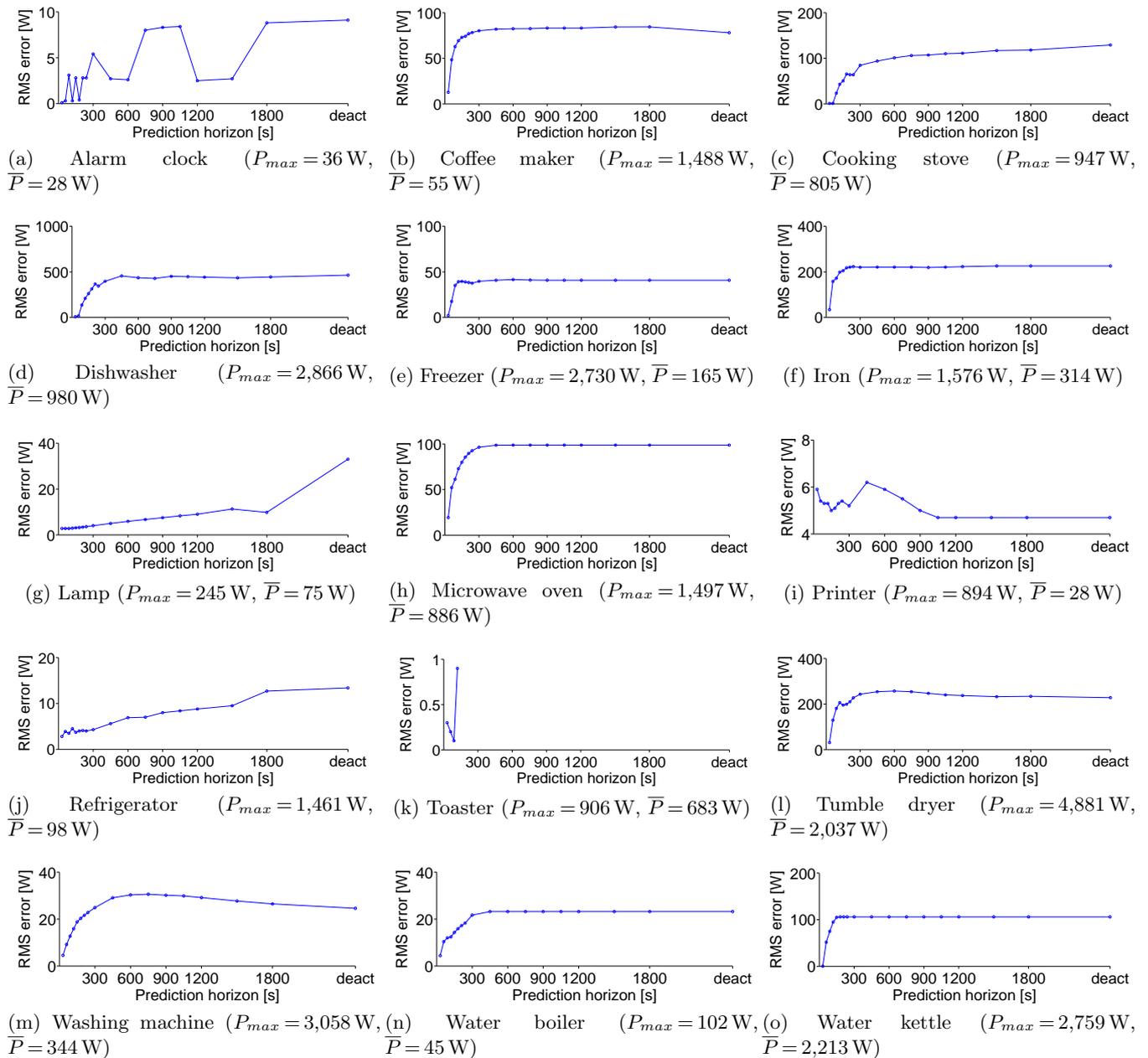
Firstly, the table conveys that more generic signatures exist for the shorter prediction horizon of just 60 seconds, as compared to the other two settings. This can be explained by referring back to Fig. 1, from which it can be visually confirmed that readings closer to the beginning of an activity segment are much more similar than in later operation phases. As a result, short-term predictions can be made more easily, which the figures in the table confirm. Secondly, it can be seen that across all appliance types, on average only 4% of the signatures are generic, whereas around 50% are ambiguous and 46% unique. While not a performance indicator of our system, but rather a characteristic of the underlying power consumption data, these figures help us assess the system’s operability on memory-constrained sensor nodes. In general, the results show that (with the exception of the toaster appliance) at least one generic signature could always be extracted.

## 5. EVALUATION

In this section, we evaluate to which our approach can allow smart plugs to emit power consumption predictions. We start with an analysis of the class-wise prediction capabilities before analyzing signature repositories that comprise signatures for all appliance types. A discussion of our findings as well as the observed limitations of our approach, including an analysis of its applicability on embedded systems, concludes this section.

### 5.1 Class-wise Prediction Model

In the first experiment, we have analyzed to which extent the extracted signatures (cf. Table 2) are suited to predict appliance deactivations. To this end, we have evaluated the accuracy of the predictions emitted by the system by means of their RMS deviation from the appliance’s actual consumption. We have considered different prediction horizons between 30 seconds to 30 minutes, as well as predicting the power consumption until the appliance’s deactivation. The resulting prediction errors are visualized in Fig. 4 for each of the 15 appliance types under consideration. All subfigures are furthermore annotated by the maximum power draw,  $P_{max}$ , as well as the mean power consumption across all activity segments ( $\bar{P}$ ) in order to put the results into perspective.

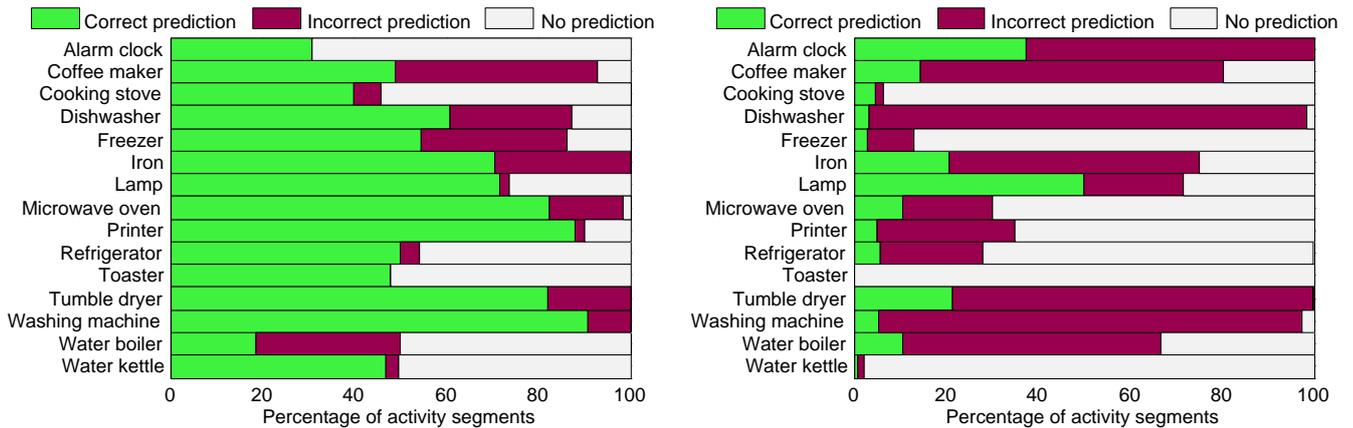


**Figure 4: Comparison of prediction errors for all considered appliances when using class-internal signature repositories with generic and unambiguous signatures.**

From the results, a trend that the prediction errors significantly depend on the type of the underlying appliance can be observed. Relatively small errors were reported for most devices for which (a) more training data was available and (b) well-defined operational cycles dominate the power consumption. This is, e.g., the case for the refrigerator, the freezer, or the water kettle. While both conditions also hold for the dishwasher appliance, however, its prediction errors are much larger and thus an unconditional generalization of aforementioned observations is not possible. Still, with the exception of the printer, all graphs indicate the excellent short-term predictability of the attached loads, while

forecasts with longer horizons generally lead to increased prediction errors.

Besides analyzing how accurate the signatures permit to predict the appliance’s future consumption, however, it is also important to determine how many signature occurrences can actually be found in real-time data. For each segment in the testing set, we have thus assessed whether it was covered by a prediction at all, and whether the prediction was within the tolerable bounds. The results are visualized in Fig. 5a and show the fraction of segments for which correct predictions have been made (green), for which incorrect forecasts have been emitted (purple), and for which no predictions



(a) Class-wise distribution of signature matches per appliance type after ambiguous signatures have been removed (120 seconds prediction horizon).

(b) Distribution of signature matches per appliance type for a common repository with generic signatures of all appliance types (120 seconds prediction horizon).

**Figure 5: Comparison between coverage of testing data segments for class-wise and combined signature repositories.**

have been made at all. On the one hand, the figures shows that for many appliances correct predictions are emitted for more than 70% of the testing traces, and only few false positive predictions are reported. On the other hand, however, the false positive rate of the solution is significant for some appliance types like the coffee maker. To sum up, our results have shown that class-wise consumption forecasting does not work perfectly across all 15 analyzed device types, but still succeeds in making forecasts with small error for some of them (e.g., the printer and the washing machine).

## 5.2 Combined Prediction Model

Next, we combine the signatures with multiple occurrences from the previous analysis into a single model. While ambiguities within each appliance class were eliminated in the previous analysis, confusion may be newly introduced due to the merging of many individual signature repositories. Hence, a second repetition of the signature quality assessment step has been applied to ensure that no ambiguities exist in the resulting signature repository. We visualize the impact of merging all individual signatures into a common repository in Fig. 5b. The figure clearly shows that significantly less correct forecasts can be made based on the combined model. In fact, our systems no longer succeeds in making more than 50% of correct predictions, while up to 90% wrong predictions result for the washing machine and the dishwasher now. Combining multiple signature repositories is thus diametrically opposed to the goal of making accurate predictions.

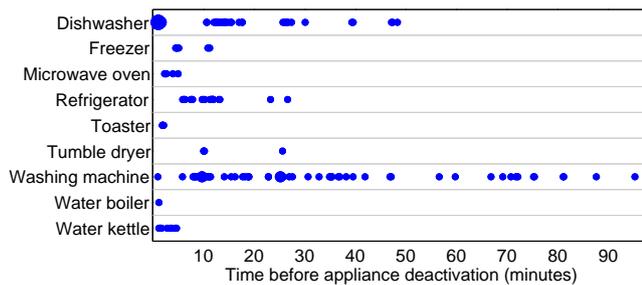
## 5.3 Observations and Discussion

The results indicate that many device classes with well defined operational cycles (like refrigerators or washing machines) show a comparably large number of unambiguous signatures when analyzed individually. This allows our approach based on time series pattern matching to emit accurate predictions of their future power demand. Given the comparably small size of the training and testing data set (for some appliances less than 20 activity segments were

available in the data), we expect an even higher ratio of correct forecasts in practical systems where a larger set of historical data is available for training.

A further insight is that significantly less predictions can be made when a combined model is being used as compared to the case when appliance types are regarded independently (i.e., comparing the results in Fig. 5). Instead of using a combined model, it might thus prove beneficial to prepend the presented system by a system that categorizes the underlying appliance type first (e.g., [21]). As most household appliances remain connected to the same wall outlet for most of their operational time, this step needs to be performed once upon system installation and only needs to be repeated when a high number of erroneous predictions are being made.

While we have conducted all prior studies on a server computer in order to analyze the full parameter space, the system's performance on embedded systems is crucial for its applicability on smart plugs and smart meters. The initial configuration (i.e., the complete trace preprocessing if historical input data and the extraction and assessment of signatures) is only required during the system's configuration, and it can be safely assumed that it will be executed on a server machine. Subsequent processing steps for the real-time data, however, need to be executed on the embedded systems, where the available amount of memory limits the number of signatures and forecasts that can be stored. For each power reading, 16 bits are required (in order to capture all values occurring in the data), such that each signature occupies 40 bytes of memory. Two more bytes per second are required to store the corresponding predictions in the node's memory, which in turn are required in order to emit the forecast. Assuming the case of when predictions shall be emitted for 10 minutes, and 77 signatures have been extracted (as observed in case of the microwave oven), a memory requirement of 1.5 kBytes for the signatures and 45 kBytes for the predictions is given. While these would



**Figure 6: Visualization of the achievable prediction horizon when using the resulting signatures (marker size indicates occurrence frequency).**

barely fit into the TelosB’s application memory of 48 kBytes size, no space would remain for the actual application implementation. A trade-off thus needs to be found between the targeted prediction horizon and the number of signatures in the repository.

In a supplementary experiment, we have quantified the maximum achievable prediction horizon when considering unambiguous and generic signatures. To this end, we have analyzed the temporal distance between a signature match and an appliance’s deactivation. Please note that this is orthogonal to the previous analysis, in which we have determined how accurate an appliance’s power demand can be detected, because we now investigate how long in advance it can be identified whilst ignoring its power consumption until then. The results are shown in Fig. 6. The figure shows that the resulting signature set permits the prediction of the remaining operation cycle of the washing machine more than 90 minutes in advance. Moreover, it confirms that the achievable prediction horizon for appliances with defined operational schedules can reach tens of minutes easily. This confirms our assumption that autonomously operating devices with predefined operation cycles emit characteristic features which can be used to estimate their remaining operation time.

As a final remark, the approach presented in this paper has been based on device-level measurements instead of using circuit-level data or even relying on meter-level data. We have deliberately chosen this approach due to the availability of a comprehensive data set which has allowed us to extract the signatures for each device individually. The resulting model can also be applied to circuit or meter level, permitted that suitable means to extract signatures from the aggregated load exist. For example, first order derivatives of the signatures could be used in order to disregard the base load offset and apply our approach to aggregated load data. While we believe that the repository of signatures extracted on per-device level can be leveraged for its application on circuit- and meter-level, the extraction of generic and unambiguous signatures from aggregate data represents a more complex task which is beyond the scope of this contribution. Recent research results (e.g., [5,10,20]) have, however, shown major breakthroughs in the capability of disaggregating load data collected at a single sampling point. These approaches are orthogonal to the prediction system presented in this paper, and can potentially be combined in order to allow for predictions of aggregate loads as well.

## 6. CONCLUSIONS

In this paper, we have analyzed to which extent the power consumption of electric appliances can be predicted based on historical consumption data. Our approach is based on the collection of consumption data from smart plugs, from which characteristic power signatures are extracted. Based on an appliance’s actual past power consumption, signatures are potential indicators for future appliance behavior. Our system hence detects the presence of these signatures in subsequently collected traces in order to forecast the remaining duration of an appliance’s operation. Our analysis was based on more than 2,700 activity segments of 15 different appliance types, and has shown that many signatures reliably indicate the device’s power demand in the near future. The evaluation has shown that in conjunction with knowledge about the underlying device type, large fractions of appliance deactivations can be correctly predicted. We believe that predictions of household loads may bring a large benefit to power grid operators. In combination with approaches to predict when appliances will be activated, a forecasting system can be implemented to match generation and demand, and thus cater for a sustainable future.

## 7. REFERENCES

- [1] H. K. Alfares and M. Nazeeruddin. Electric Load Forecasting: Literature Survey and Classification of Methods. *International Journal of Systems Science*, 33(1):23–34, 2002.
- [2] S. Aman, Y. Simmhan, and V. K. Prasanna. Improving Energy Use Forecast for Campus Micro-grids using Indirect Indicators. In *Workshop Proceedings of the 11th International Conference on Data Mining (ICDM)*, pages 389–397, 2011.
- [3] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Bergés. BLUED: A Fully Labeled Public Dataset for Event-Based Non-Intrusive Load Monitoring Research. In *Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, pages 1–5, 2012.
- [4] O. Ardakanian, S. Keshav, and C. Rosenberg. Markovian Models for Home Electricity Consumption. In *Proceedings of the 2nd Green Networking Workshop (GreenNet)*, pages 31–36, 2011.
- [5] N. Batra, J. Kelly, O. Parson, H. Dutta, W. Knottenbelt, A. Rogers, A. Singh, and M. Srivastava. NILMTK: An Open Source Toolkit for Non-intrusive Load Monitoring. In *Proceedings of the 5th International Conference on Future Energy Systems (ACM e-Energy)*, pages 265–276, 2014.
- [6] M. Beccali, M. Cellura, V. Lo Brano, and A. Marvuglia. Short-Term Prediction of Household Electricity Consumption: Assessing Weather Sensitivity in a Mediterranean Area. *Renewable & Sustainable Energy Reviews*, 12(8):2040–2065, 2008.
- [7] J. Bradley, J. Barbier, and D. Handler. Embracing the Internet of Everything to Capture Your Share of \$14.4 Trillion. Cisco White Paper, [http://www.cisco.com/web/about/ac79/docs/innov/IoE\\_Economy.pdf](http://www.cisco.com/web/about/ac79/docs/innov/IoE_Economy.pdf), 2013.
- [8] C. Chen, B. Das, and D. J. Cook. Energy Prediction Based on Resident’s Activity. In *Proceedings of the 4th International Workshop on Knowledge Discovery from Sensor Data (SensorKDD)*, pages 1–7, 2010.

- [9] B. Dong, C. Cao, and S. E. Lee. Applying Support Vector Machines to Predict Building Energy Consumption in Tropical Region. *Energy and Buildings*, 37:545–553, 2005.
- [10] D. Egarter, V. P. Bhuvana, and W. Elmenreich. PALDi: Online Load Disaggregation via Particle Filtering. *IEEE Transactions on Instrumentation and Measurement*, 2014. accepted for publication.
- [11] ELEXON Ltd. Load Profiles and Their Use in Electricity Settlement. Online: [http://www.elexon.co.uk/wp-content/uploads/2013/11/load\\_profiles\\_v2.0\\_cgi.pdf](http://www.elexon.co.uk/wp-content/uploads/2013/11/load_profiles_v2.0_cgi.pdf), 2013. Last access on 13 February 2014.
- [12] Ericsson. More than 50 Billion Connected Devices. Ericsson White Paper 284 23-3149 Uen, 2011.
- [13] D. Evans. The Internet of Things – How the Next Evolution of the Internet is Changing Everything. Cisco Internet Business Solutions Group (IBSG) White Paper, 2011.
- [14] D. Gantenbein, C. Binding, B. Jansen, A. Mishra, and O. Sundström. EcoGrid EU: An Efficient ICT Approach for a Sustainable Power System. In *Proceedings of the 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT)*, pages 1–6, 2012.
- [15] International Energy Agency. Technology Roadmap Smart Grids. Online: [http://www.iea.org/publications/freepublications/publication/smartgrids\\_roadmap.pdf](http://www.iea.org/publications/freepublications/publication/smartgrids_roadmap.pdf), 2011. Last access on 13 February 2014.
- [16] J. Z. Kolter and J. Ferreira Jr. A Large-Scale Study on Predicting and Contextualizing Building Energy Usage. In *Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1349–1356, 2011.
- [17] J. Z. Kolter and M. J. Johnson. REDD: A Public Data Set for Energy Disaggregation Research. In *Proceedings of the 1st Workshop on Data Mining Applications in Sustainability (SustKDD)*, pages 1–6, 2011.
- [18] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajic. AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research. In *Proceedings of the Electrical Power and Energy Conference (EPEC)*, pages 1–6, 2013.
- [19] O. Monnier. A Smarter Grid with the Internet of Things. Texas Instruments White Paper, <http://www.ti.com/lit/ml/slyb214/slyb214.pdf>.
- [20] O. Parson, S. Ghosh, M. Weal, and A. Rogers. An Unsupervised Training Method for Non-intrusive Appliance Load Monitoring. *Artificial Intelligence*, 2014. accepted for publication.
- [21] 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 *Proceedings of the 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT)*, pages 1–9, 2012.
- [22] The Telegraph. Scottish wind farms paid £1 million to shut down one day. Online: <http://bit.ly/10cWBjF>, 05 May 2013. Last access on 13 February 2014.
- [23] N. C. Truong, L. Tran-Thanh, E. Costanza, and S. D. Ramchurn. Activity Prediction for Agent-based Home Energy Management. In *Proceedings of the 4th International Workshop on Agent Technologies for Energy Systems (ATES)*, pages 1–8, 2013.
- [24] H. L. Willis. *Spatial Electric Load Forecasting*. CRC Press, 2nd illustrated and revised edition, 2002.
- [25] J. Yang, H. Rivard, and R. Zmeureanu. On-line Building Energy Prediction using Adaptive Artificial Neural Networks. *Energy and Buildings*, 37:1250–1259, 2005.
- [26] Z. Yua, F. Haghghat, B. C. Fung, and H. Yoshino. A Decision Tree Method for Building Energy Demand Modeling. *Energy and Buildings*, 42:1637–1646, 2010.