The Potential of Smart Home Sensors in Forecasting Household Electricity Demand

Holger Ziekow  
AGT International  
Darmstadt, Germany  
hziekow@agtinternational.com  

Jens Strüker  
Fresenius University of Applied Sciences  
Idstein, Germany  
jens.strueker@hs-fresenius.de  

Christoph Goebel  
Technical University Munich  
Munich, Germany  
christoph.goebel@tum.de  

Hans-Arno Jacobsen  
Technical University Munich  
Munich, Germany  
jacobsen@in.tum.de  

Abstract—The aim of this paper is to quantify the impact of disaggregated electric power measurements on the accuracy of household demand forecasts. Demand forecasting on the household level is regarded as an essential mechanism for matching distributed power generation and demand in smart power grids. We use state-of-the-art forecasting tools, in particular support vector machines and neural networks, to evaluate the use of disaggregated smart home sensor data for household-level demand forecasting. Our investigation leverages high resolution data from 3 private households collected over 30 days. Our key results are as follows: First, by comparing the accuracy of the machine learning based forecasts with a persistence forecast we show that advanced forecasting methods already yield better forecasts, even when carried out on aggregated household consumption data that could be obtained from smart meters (1-7%). Second, our comparison of forecasts based on disaggregated data from smart home sensors with the persistence and smart meter benchmarks reveals further forecast improvements (4-33%). Third, our sensitivity analysis with respect to the time resolution of data shows that more data only improves forecasting accuracy up to a certain point. Thus, having more sensors appears to be more valuable than increasing the time resolution of measurements.

Keywords—Smart Grid; Smart Home; Forecasting; Value of ICT

I. INTRODUCTION

In accordance with the so-called 20-20-20 goals, the European member states have to obtain at least 20 percent of their electricity from renewable sources by 2020. Japan recently decided to raise the share of renewable power to more than 30 percent of its energy mix over the next three decades and Germany even strives for 80 percent by 2050. In the long term, the integration of a high share of renewables is going to transform the entire electricity system. The fastest growing renewable resources, wind and solar, are highly variable; power generation from these resources is not dispatchable (cannot be controlled on demand), is intermittent (exhibits large fluctuations), and is uncertain (random or not known in advance). The operation of today’s power grids is centered on controlling dispatchable generators in ways such that demand is satisfied at all times. The integration of renewable energy resources into the power grid makes this task increasingly difficult and calls for more balancing power that can be activated upon short notice to maintain grid stability [1].

An alternative to fast ramping dispatchable power plants for providing balancing power is – to a certain degree – demand side management (DSM). Recent government reports view demand side management as a promising way to cope with the decreasing stability of power grids and the increasing price risks in spot markets for retailers (e.g. [11]).

As of today, demand side management is practiced primarily by large industrial customers (so-called demand response). Only very rarely it is applied by private households (e.g. [6]). However, things are changing rapidly. First, the numbers of installed distributed energy resources have sharply increased: In 2012, the share of renewables has crossed the mark of 20 percent in Germany. Accordingly, the relevance of residential load-shifting has, at least in relative terms, increased. Second, with more and more smart meter roll-outs, home automation applications and almost ubiquitous Internet access, we are witnessing a shift to smarter power distribution grids. Increasing amounts of load will be dispatched by smart grid systems to achieve certain targets, such as maximal local consumption of solar power, while staying with the limitations imposed by the installed power grid equipment, e.g. transformers and power lines. A basic precondition for such systems to operate effectively is for them to be aware of dispatchable and non-dispatchable demand at different locations in the distribution grid. Forecasting demand at this highly disaggregated level and at the small time scales necessary for distribution grid control is therefore a highly relevant and challenging problem. Despite this fact, there is still little published work on household-level electricity demand forecasting. A particularly relevant question from the perspective of ICT management is to research the impact of...
higher data granularity on the accuracy of such demand forecasts: For the sake of realizing smart grid systems at minimal financial cost, it would be highly useful to characterize the trade-off between the benefits of higher forecasting accuracy and the cost of deploying more hardware and using more computational time to obtain and process the necessary data. Due to the currently still low adoption of home automation applications it is difficult to obtain real-world data that could help us to understand how more disaggregated measurements, both in the device and temporal dimension, could contribute to household-level demand forecasting. The contributions of this paper can be seen as a first step in this direction.

This paper is organized as follows. In Section II we review relevant literature and identify a specific lack in research. Section III describes our experimental setup for analyzing the impact of smart home sensors on household demand forecasting and Section IV presents the results of our experiments. In Section V we discuss our findings and provide an outlook to future research.

II. RELATED WORK AND LACK OF RESEARCH

The demand side of electricity system receives increasing attention by research and industry. At the same time, demand side management is still exceptional in the residential sector of electricity markets (e.g. [4]). Utilities such as Nevada utility NV Energy have just begun to roll out demand response programs that use home energy gadgets and smart thermostats in consumers’ homes (www.nvenergy.com/mpowered). The lack of knowledge about the extent of the potential benefits is said to be one of the contributory factors for current status of residential demand response applications [4]. The benefit of improved consumption forecast in itself is part of this bigger lack of knowledge and can be divided according to the different demand side management providers: distribution grid operator, electricity retailer and energy management provider.

Demand side management (DSM) is a portfolio of measures to improve the energy system at the side of consumption [27]. A distribution grid operator utilizes improved consumption forecasts in order to stabilize grids with a high share of distributed renewables. The objective is to more effectively balance supply and demand in the power grid and, in this way, reduce the need for additional operating reserves (e.g. [9], [21], [14]). The integration of intermittent resources of power generation, such as wind and solar, will lead to unprecedented energy price fluctuations and increasing price uncertainty [13]. Electricity retailers can meet this challenge by partly adapting the demand side of the electricity markets to the fluctuating supply side [13]. Thus, improved consumption forecasts provide the retailer with, for instance, a strategy to select the most suitable customers for implementing demand response programs. A third group benefitting from better forecast are residential energy management providers. Using an Internet-enabled thermostat, companies such as EcoFactor or NEST gather and analyze a specific heating and cooling system information, thermostat settings, personal preferences, indoor temperatures, local weather conditions and several other behavioral factors. These services then continually learn and adapt to these factors by making small adjustments to thermostat settings throughout the day to reduce energy consumption and shift energy use without compromising comfort (e.g. www.ecofactor.com, [8]).

Apart from a number of contributions on the load shifting potential of residential customers (e.g. [32]), recent research has been done in particular on the impact of dynamic tariffs and the communication of price signals to customers (e.g. [17], [2]). The advent of new consumer-facing technologies and local energy management systems that communicate with the electric utility or third party energy management providers (e.g. ecotope or opower) currently shifts the interest to automated DSM. Accordingly, an integrated optimization of financial savings in case of multi-tariff contracts and DSM is increasingly discussed and researched ([30], [18], [15]). While demand response services that engage consumers are an important emerging aspect of the smart grid, it comes as a surprise that the underlying cause-effect-chain between disaggregated power measurements and demand forecast accuracy has so far been neglected. There are only few works dealing with this important aspect (e.g. [24], [36], [16]). Other researchers have looked at smart household appliances, but merely on an aggregated level (e.g. [20]), without real data on household level and take the impact of fine-granular consumption data on consumption forecasts for granted. In the context of profiling, research on data-privacy in smart grids extensively considers demand forecasting based on more fine-granular billing and consumption data. However, today’s most popular solutions propose minimizing, anonymizing or concealing consumption data (e.g. [10], [5]) instead of alternatively giving consumers control over their data without altering data quality [34]. We can thus state, that the impact of disaggregated power measurements on demand forecast accuracy currently lacks research.

III. EXPERIMENTAL SETUP

In this section, we describe the experimental setup that we have used to evaluate the impact of disaggregated appliance-level demand data on the household-level demand forecast. It covers the technical solution for collecting consumption data as well as information about the selected households. Furthermore, we describe the implementation of the evaluated forecasting methods and the necessary data pre-processing.

A. Data Collection

We collected consumption data from private households as part of a field trial for smart building technology. This trial is executed as part of the pilot project “PeerEnergyCloud” that addresses load balancing in the distribution grid.

For the data collection we used wireless ZigBee sensors that were installed between power sockets and electrical devices in the private households of volunteers. The sensors measure energy consumption in kWh and wirelessly transmit their latest measurements to a data collection gateway once every 2 seconds. From the gateway, the data gets forwarded to a central database for analysis. An overview of the setup is given in Fig. 1 and details about the technical solution can be found in [35].
We randomly selected three family households in a small German city from the list of volunteers. All households were equipped with 5-7 plug meters measuring the energy consumption of the different devices. Table 1 lists the monitored devices in each of the selected households.

<table>
<thead>
<tr>
<th>Household</th>
<th>Monitored Devices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household 1</td>
<td>computer, fridge, two TVs, stereo, freezer</td>
</tr>
<tr>
<td>Household 2</td>
<td>freezer, two fridges, lamp, dryer, TV, washing machine</td>
</tr>
<tr>
<td>Household 3</td>
<td>dryer, two fridges, TV, washing machine</td>
</tr>
</tbody>
</table>

For the analysis, we collected data over 30 days. Fig. 2 provides insights with respect to the characteristics of this data. All figures show the average power demand computed over time intervals of 15 minutes. Fig. 2 (top) shows the detailed power consumption of household 1 on a randomly chosen day during the observation period. The breakdown in device-specific consumption patterns shows how different appliances drive the total household demand. Fig. 2 (bottom) depicts the total demand of all measured devices for each household over the course of one week. The charts reveal large variations between the households and different days. The sharp spikes in the curves for household 2 and 3 stem from the dryers and washing machines, which also cause higher peak loads compared to household 1. The high variability of household demand underlines the challenges involved in computing accurate consumption forecasts for individual households.

B. Forecasting Methods

Our work focuses on the impact of the provisioned data granularity on forecasting, not the forecasting method itself. That is, our aim was not to develop a dedicated forecasting method. Instead, we deliberately relied on standard forecasting tools. To obtain the results presented in this paper, we leveraged the time-series-analysis mechanism of the well-known machine learning software Weka (version 3.7.7) (WEKA, 2012) with the default forecasting model and default settings. The default forecasting mechanism in Weka is a support vector machine (SVM) [31]. However, to exclude our results being an artifact of the use of SVM-based forecasting, we repeated the experiments using artificial neural networks (ANN) as forecasting mechanism.

C. Experiment Design

The goal of the experiments was to test how the granularity of consumption measurements impacts the accuracy of demand forecasts. In particular, we investigated the impact of two factors: (i) the temporal resolution of the data, and, (ii), the effect of aggregated measurements, as done by smart meters, versus device-specific measurements, as done by smart home infrastructures. Therefore we used the sensor data from the field trial to create time series of consumption measurements with different level of detail.
The raw sensor data were collected with sampling intervals of two seconds, where each value reports the total accumulated energy consumption of the monitored device in kWh. From these measurements we derived the energy consumption in a given time slice by taking the latest value in that slice and subtracting the last value from the previous slice. The sensor measurements had some gaps due to network errors. We therefore set time slices with no measurements to zero. We subsequently denote the derived consumption of device D in a time slice of length \( t \) as \( C(D)_t \). The summed up consumption of all devices in a household in a time slice is denoted as \( C(\text{total})_t \).

For the experiments we constructed input vectors \( V_i \) with data of different granularity and compared the accuracy of the corresponding forecasts. Specifically, we analyzed two dimensions of granularity. The first investigated dimension was the breakdown of consumption measurements to individual electrical devices. Therefore we (i) constructed input vectors with the sum of the device specific measurements, and, (ii), constructed vectors that additionally included the consumption of each device. The second investigated dimension was the temporal resolution of sensor measurements. Therefore we constructed vectors with differently sized time slices. Input vectors always also included the respective lower resolution values to make sure that the learning algorithm had at least the same information at its disposal as in the lower resolution case. Table 2 provides an overview of how we constructed input vectors for the different experiments.

Table 2 Input Parameters for the Experiment

<table>
<thead>
<tr>
<th>Time Series Name</th>
<th>Input Vectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment Set 1: Impact of Device Specific Measurements (15 min time slices)</td>
<td>( V = {t, C(\text{total})_{t+1}, t \in T} )</td>
</tr>
<tr>
<td>device specific consumption</td>
<td>( V = {t, C(D)_t, \ldots, C(DN)_t, t \in T} ) where ( D_1, \ldots, DN ) are the devices in the respective households.</td>
</tr>
<tr>
<td>Experiment Set 2a: Impact of Temporal Resolution (15 min time slices)</td>
<td>15 min sample intervals: ( V = {t, C(\text{total})_{t+1}, t \in T} )</td>
</tr>
<tr>
<td>5 min sample intervals: ( V = {t, C(\text{total})<em>{t+1}, C(\text{total})</em>{t+2}, \ldots, C(\text{total})_{t+5}, t \in T} )</td>
<td></td>
</tr>
<tr>
<td>1 min sample intervals: ( V = {t, C(\text{total})<em>{t+1}, C(\text{total})</em>{t+2}, \ldots, C(\text{total})_{t+15}, t \in T} )</td>
<td></td>
</tr>
<tr>
<td>Experiment Set 2b: Impact of Temporal Resolution (1h time slices)</td>
<td>60 min sample interval: ( V = {t, C(\text{total})_{t+1}, t \in T} )</td>
</tr>
<tr>
<td>1 min sample interval: ( V = {t, C(\text{total})<em>{t+1}, C(\text{total})</em>{t+2}, \ldots, C(\text{total})<em>{t+15}, C(\text{total})</em>{t+16}, \ldots, C(\text{total})_{t+30}, t \in T} )</td>
<td></td>
</tr>
</tbody>
</table>

IV. Results

Our experimental results reveal how access to more detailed consumption measurements impact the quality of energy consumption forecasts. This section describes these results in detail.

A. Impact of Device Specific Measurements

We conducted the first set of experiments to analyze the impact of device specific measurements as opposed to measurements of the total consumption of a household. As measure for the prediction accuracy we use the mean absolute error (MAE).

To measure this effect, we compared the prediction accuracy of models fitted to time series of (i) “device specific consumption” and, (ii), of “total consumption” at each data point (see experiment set 1 in Table 2). We thereby approximated a comparison of demand forecasting based on smart metering data and smart home sensor data.

The results are displayed in Fig. 3. For illustration the top part of the figure provides sample plots of predicted load, the actual load and the four weeks average. The bottom part of the figure shows the error achievement by forecasting based on device specific measurements compared to forecasting using aggregated consumption. The accuracy gains from forecasting based on aggregated to forecasting based on disaggregated data very strongly between the investigated households. However, in all three cases we observe a significant gain of the forecasting accuracy. In the case of household 3, we even observe a gain of more than 30%.

B. Impact of Temporal Resolution

To investigate the effect of temporal resolution, we pre-aggregated the data from the field trial to different resolution levels (see experiment set 2 in Table 2). The results for predictions of the next 15 min time slice are given in Fig. 4. The figure shows the relative error reduction of predictions with 5 min sampling intervals \(|\text{MAE}_{15} – \text{MAE}_{5}|/\text{MAE}_{15}\) and 1 min intervals compared to predictions based on 15 min intervals \(|\text{MAE}_{15} – \text{MAE}_{1}|/\text{MAE}_{15}\).

Our results indicate that a higher time resolution leads a significant improvement of the prediction accuracy across all three households. The results also show that sampling once per minute clearly improves the prediction compared to 5 min intervals in two out of three cases.
The results for predictions of the next 60-min time slice are shown in Fig. 5. Again, we compare the MAE of the prediction based on different sampling intervals with the MAE for the biggest sampling interval, in this case 60 minutes. Our results reveal that whereas sampling intervals of 15 min and 30 min have a positive effect on the 60-minute average forecast, a further increase of the sampling rate in this case has a negative impact across all households. This result can be explained by the fact that inputs of low relevance can harm the process of training the forecasting model. Together with the results from Fig. 4, our results show that in the observed cases high sample rates are not useful for longer-term forecasts (60 min) but are beneficial in short-term forecasts (15 min).

C. Comparison with a Baseline Solution

The results provided above only show the relative improvements of predictions based on different input data. While this is the core of our investigation, we ran additional tests to verify the general quality of the predictions. We deliberately did not optimize the used prediction method but used standard mechanisms in default configurations. The evaluation of the absolute prediction accuracy thus mainly serves as validation of the general approach of using machine learning and smart building sensors for demand forecasting. Furthermore, the results provide a lower bound for what we can achieve with optimized models.

In order to evaluate the overall performance, we compared the prediction results with a baseline approach. As baseline prediction we used a persistence estimator, which is known as a hard to beat for short-term forecasts of highly variable time series. Table 3 shows the results from the experiments. For all investigated households we observed improved prediction accuracy with the investigated approaches compared to the baseline, with very high gains in two out of three cases.

D. Experiments with Neural Networks

The results presented so far were obtained using a support vector machine (SVM) based forecasting method. To exclude that our results are an artifact of SVNs, we repeated the experiments using artificial neural networks (ANNs) as forecasting method.

We used the Weka implementation of a multilayer perceptron as implementation of the ANN. In a first set of experiments we again used the default setting of the Weka framework. This resulted in an extremely poor forecasting accuracy. We therefore tuned the parameters of the learning algorithm (i.e., we enabled decay in the learning and set the validation threshold 50). With the tuned parameters, the ANNs achieved similar results as the SVM-based models in most cases. However, in some test runs the ANNs still performed poorly, specifically in experiments with very fine-grained data (i.e., predictions had relative absolute errors above 100%). Altogether, the prediction error of ANNs and SVM deviated by less than 5% in 10 of 30 test runs. In 15 out of the remaining 20 cases, SVMs achieved better forecasting accuracy.

Overall, our experiments with ANNs show more variance in the results but still show the same effects that we observed in SVM-based forecasting. That is, both approaches show gains when using device-specific measurements. Also, we observe benefits for moderate increase of the sample rates but see negative effects for very fine-grained sampling.

V. DISCUSSION AND OUTLOOK

The dynamics of residential electricity demand are becoming increasingly important in context of renewable integration and demand side management. In this work, we have investigated the impact of using advanced forecasting methods and historic consumption data at different levels of granularity to predict household power consumption.

Instead of focusing on the algorithmic details of different forecasting methods, we deliberately used state-of-the-art machine learning tools without specific tuning to produce forecasts. Our results provide insights into the potential of advanced forecasting methods in this context and how standard tools would benefit from the richness of data provided by the progressing deployment of sensors in households. Based on high-resolution data from 3 private households collected over 30 days we were able to show that advanced forecasting methods already yield better forecasts (1-7% improvement with respect to benchmark) when carried out based on aggregated household consumption data that could be obtained.
from smart meters. Moreover, we found further forecast improvements (4-33%) when applying these methods to disaggregated data from smart home sensors compared to smart meter benchmarks. Finally, we can see from the results that a higher temporal resolution only makes sense up to certain degree depending on the forecasting target: To forecast 15-minute average demand, it makes sense to use data in 1-minute resolution. However, for the 1-hour demand forecast, lower sample intervals than 15 did not lead to improvements.

An interesting way to extend this research would be to analyze the required computational resources to make forecasts for large numbers of households in detail. The amount of disaggregated data and the complexity of forecasting models call for big data technologies to leverage the accuracy gains shown in this paper for mass deployment. However, since single forecasting tasks (i.e., one household at a certain time) can be executed in parallel, we believe that systems for disaggregated demand forecasting could be designed for high scalability.

ACKNOWLEDGEMENTS

This work has been supported in part by the German Federal Ministry of Economics and Technology, project PeerEnergyCloud which is part of the Trusted Cloud Program.

REFERENCES