Disaggregation of household load profiles

Michael Hinterstocker[1], Paul Schott, Serafin von Roon

Forschungsgesellschaft für Energiewirtschaft, Am Blütenanger 71, 80995 München,
+49 89 158121 53, mhinterstocker@ffe.de, www.ffegmbh.de

Abstract:

The ongoing digitalization enables new opportunities for research concerning the energy sector. Smart meters are utilized to measure residential energy consumption. This temporally highly resolved data can be used for investigating the energy consumption behavior of households in detail. Based only on the data generated by smart meters, a method which identifies common household appliances is developed. In contrast to other approaches which gather this information via surveys or direct measurements, the presented procedure allows to analyze a large amount of households. The appliance's existence and further details, like program duration, energy consumption per use and operation times are identified. Based on these results, the temporal behavior can be determined which allows conclusions about daily and seasonal characteristics. These findings enable detailed investigations of the Demand Side Management potential of households.

Keywords: Disaggregation, Demand Side Management, Smart Meter

1 Introduction

The continuous expansion of renewable energy production capacities like wind and photovoltaics (PV) causes challenges for the system. These energy sources are characterized by a volatile electricity production which is mainly dependent on the current weather conditions. In contrast to power plants, their generation cannot be planned, but only approximated by forecasts. In order to keep balance between generation and demand, compensatory measures are required. In this context, Demand Side Management (DSM) is an option to adjust energy consumption to generation. In private households some appliances are suitable for postponing the operation: dishwasher, washing machine, dryer, refrigerator and freezer. To investigate the exact DSM potential some information about the appliances is necessary:

- How many appliances are present in a household?
- How often are the appliances used?
- Which characteristics do the load profiles have?
- What is the energy consumption per use?
- How long does a program last?
- When are the appliances used?

To gather this information, three different approaches are possible: a survey, measuring the appliances directly or the disaggregation of the household's power consumption. Due to some disadvantages of the first two listed procedures (e. g. survey: inaccuracy, measuring: great effort) this paper focuses on disaggregation – Figure 1 illustrates the concept. The first plot contains the total

[1] Jungautor

household load, whereas the second plot visualizes the load profiles of single appliances which were used during the displayed time period. As already mentioned, the challenge is to implement this kind of algorithm for a variety of households. With respect to the spread of smart metering, the required data is getting collected for many households. The available data basis consists of around 2 500 German households. Due to partially incomplete recording, these are reduced to 565 with good data quality from August 2014 to July 2015 for further analyses. [1]

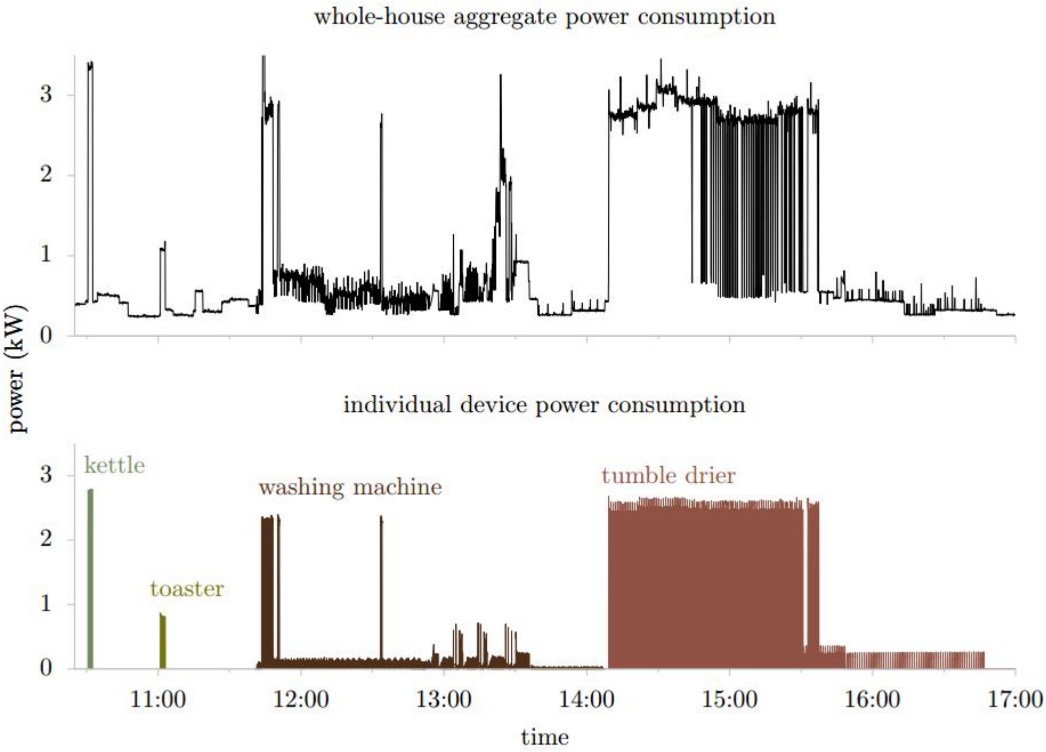

Figure 1: *Example for disaggregation [2]*

2 Methods

Existing approaches rely on a known load profile of the appliance to identify. Therefore, all different appliances' programs have to be measured. Based on these data, algorithms were implemented which can recognize this load profile. This approach is impractical for a lot of households due to the effort which has to be spent on measuring every single appliance and program. [2, 3]

For this reason, the goal of this work is to design an algorithm which can disaggregate the power consumption without prior knowledge about the present appliances. The method consists of two sections, which are applied separately for every household: Recognizing of load profiles which occur repeatedly and subsequently assigning them to an appliance depending on defined properties. Dishwasher, washing machine and dryer are considered here due to their comparably high DSM potential and user acceptance.

2.1 Identification of dishwashers and washing machines

In Figure 2, seven steps are illustrated which are applied to disaggregate the total power consumption. The first six steps identify regularly appearing load profiles. The last step includes the assignment of the determined load profiles to appliance types. In the end, the result is none, one or more load profiles per appliance which represent different programs. In detail this is reached by:

- Extraction of relevant intervals: In a first step, the whole year's data is reduced by extracting intervals from the total power consumption which potentially are household appliances. Therefore, the load profile needs to exceed both a defined minimal power and minimal duration after the base load has been subtracted from the household's load profile. In order to take the alternating load profiles especially from dryers into account an additional time is specified. This variable defines how long the load profile may be lower than the minimal power so that the interval is still extracted as one coherent sequence.

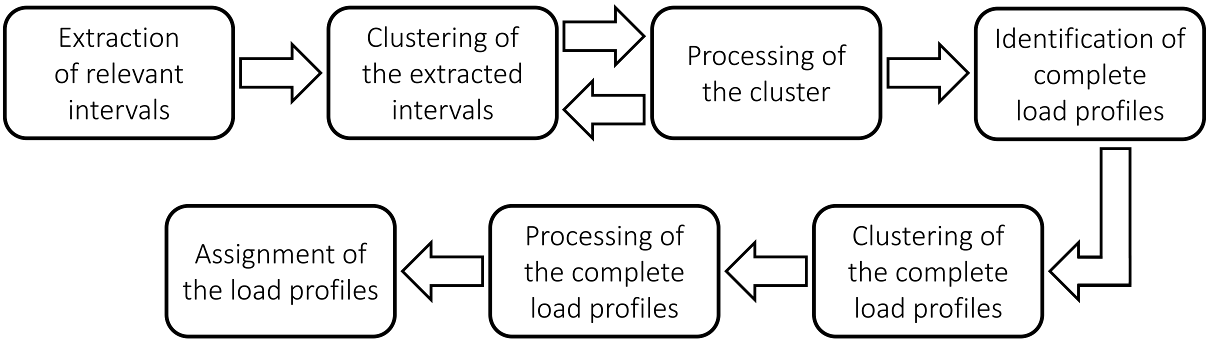

Figure 2: *Method for disaggregation*

- Clustering of the extracted intervals*:* Similar load profiles are grouped by clustering all extracted intervals. The correlation coefficient is applied as metric for clustering the load profiles. Due to high computational effort, this process is limited to three months (November–January). This has no influence on the analyses of actual uses. The result of this step is a certain amount of clusters which contain a minimum number of similar load profiles. Clusters which don't reach the minimum limit of included patterns are deleted, because it is assumed that the three considered household appliances are used regularly.

- Processing of the cluster*:* After clustering the intervals they exhibit different lengths due to possible disturbances like simultaneously running appliances. During this step, unnecessary information is deleted. The clustering step as well as the processing is repeated twice in order to improve the quality of the results.

- Identification of complete load profiles*:* The first three steps may lead to partially incomplete load profiles. It's possible that some appliances contain a heating phase which occurs later than the distance defined for extracting the intervals. Therefore, the determined load profiles are used to identify similar sections during the full year's load data. As soon as an interval with the same length exceeds the selected correlation coefficient, an interval with a fixed length (here: 150 min) is extracted. It may be expected that the maximum program duration of the three considered appliances amounts to approximately 150 min.

- Clustering of the complete load profiles: All extracted intervals of the previous step are clustered again. It appears that different programs start with the same energy consumption and the rest of the complete profiles vary.

- Processing of the complete load profiles: This step contains the same procedure like the first processing step but now the clusters of the whole year are edited.

- Assignment of the load profiles*:* At this point, several load profiles exist for every household which occur regularly. To finish the disaggregation, it is necessary to assign these load profiles to one of the three household appliance types. For this purpose, a few known appliances' load profiles are available. A conceivable brute-force-approach would be possible by looking for matches between measured and extracted load profiles. Due to a small data basis and a lot of households, it is expected that this approach leads to inadequate results.

Therefore, the load profiles are used to realize a more general approach. Five properties (depicted in Figure 3) are defined which describe the load profiles' characteristics in a more abstract way and are used to assign the determined load profiles to the appliance types:

- Number of peaks: The number of peaks or rather heating phases included in the load profile (green rectangles)
- Power increase: The amount the power increases at the beginning of a peak (blue arrow)
- Power decrease: The amount the power decreases at the end of a peak (red arrow)
- Relation of peak duration and whole duration: The sum of the duration of all heating phases divided by the whole length of the profile
- Longest peak's duration: For this load profile, the first peak is simultaneously the longest (yellow arrow).

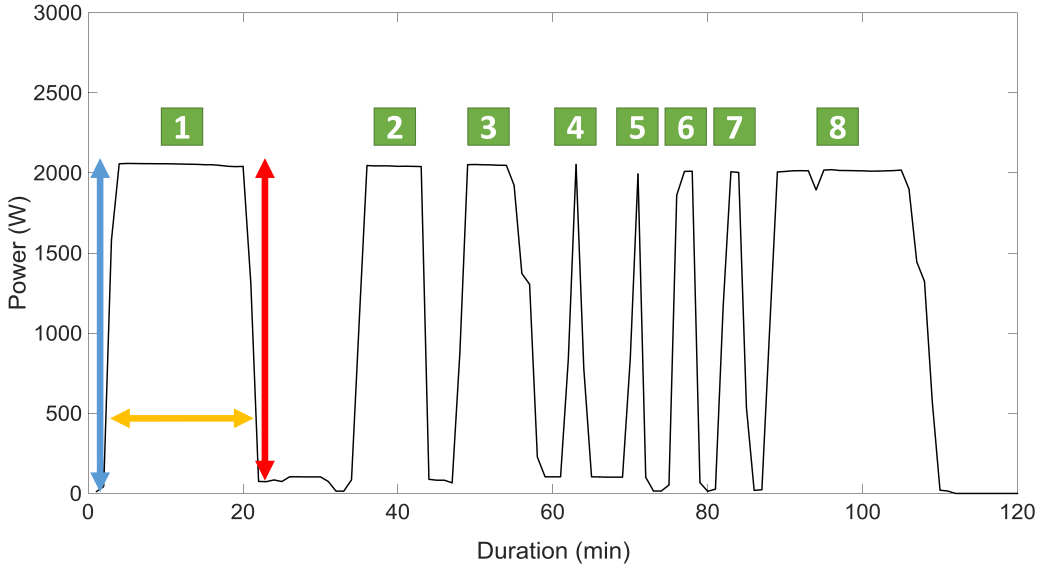

Figure 3: *Visualization of defined properties*

These five features are computed for the measured household appliances. Due to differences of the individual appliances, this yields an interval per appliance type for every property. For preventing double or triple assignment, at least one property of two appliances must be disjoint. If a load profile does not fulfill all properties for one appliance, it is neglected.

Analyses show that this approach does not work for dryers. For the 565 households, nearly no load profile is identified as a dryer. For less than 5 % of the households a dryer is assigned, which is significantly smaller than the value of the federal statistical office (39.1 %) [4]. Detailed investigation of the dryers' load profile characteristics shows why a separate procedure is necessary to identify those appliances (see chapter 2.3).

2.2 Determination of dishwasher and washing machines uses

With the method described above none, one or more than one load profiles are assigned to an appliance for every household. The load curve for the whole year is analyzed for uses by comparing the correlation coefficient between the load profile and the interval of the energy consumption. In contrast to the cluster algorithm, the correlation coefficient is smaller to also identify disturbed appliance uses. Accordingly, the difference of the two intervals' energy consumption is also considered to ensure reliable assignments.

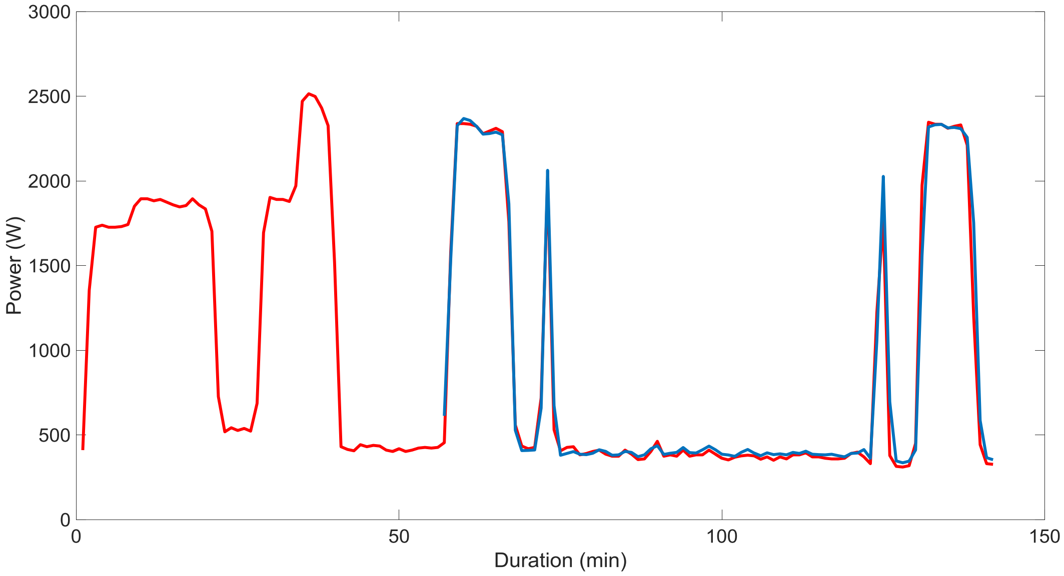

Figure 4: *Two assigned dishwasher load profiles*

As soon as more than one load profile is assigned to one appliance, two operations are required to ensure correct results. Figure 4 shows two load profiles which fulfill all properties of a dishwasher, where the blue profile is exactly a part of the red one. For analyzing the frequency of uses the following steps are applied:

- Partly deletion of the total energy consumption: If the whole year was analyzed with the blue and red load profile (see Figure 4), the frequency of uses would be too high. Every time the red load profile is recognized, the same happens for the blue one. In order to avoid duplicate uses, the appliance load profile is subtracted from the load curve, so after a match has been detected there cannot be another match for the blue profile at the same time step.
- Sorting the load profiles by their length: This step is needed to identify the correct amount of consumed energy. Searching for the shorter profile first would prevent any matches of the longer profile afterwards, and therefore underestimate the energy and duration of operation. Therefore, the load profiles have to be sorted by length in descending order.

2.3 Identification of dryers

As mentioned before, the described method does not work for dryers because nearly no load profile is recognized. European dryers are characterized by a strong alternating load profile. As soon as there is something like a measurement inaccuracy or shift the load profiles of two dryers are exactly opposite. Consequently, the correlation coefficient gets rather small or even negative so these intervals are not clustered.

Therefore, a separate method is required to identify dryers. Two additional assumptions are used here in order to reduce the solution space: Firstly, a dryer only exists if the household owns a washing machine as well and secondly, the dryer is used subsequently to the washing machine. According to the second criterion, only 4 hour intervals after washing machines are considered for the identification of dryers. In order to avoid the problem with inaccurate matches, the load curve is smoothed by converting to a time resolution of 3 minutes. After that, the same processing steps as described before can be applied and yield plausible results.

3 Results

With the described methodology of pattern identification and matching, the introduced questions can be answered.

3.1 How many appliances are present in a household?

First of all, the amount of households with a certain type of appliance is evaluated. The green bars in Figure 5 represent the number of households for which the respective appliance has been assigned. By restricting the identification method to plausible values (e. g. for dishwashers between 52 (once per week) and 730 (twice per day), this number is slightly reduced and results in the blue bars. The values of the statistical federal office are shown with orange bars for comparison [4].

The determined values for the equipment rate are smaller than the statistical values, but in the same range. This suggests that the identification method generally yields valid results, but not all occurrences are recognized. This might result from the rather small data basis of appliance load profiles. For all following analyses, only appliances that fulfill the described restrictions are considered.

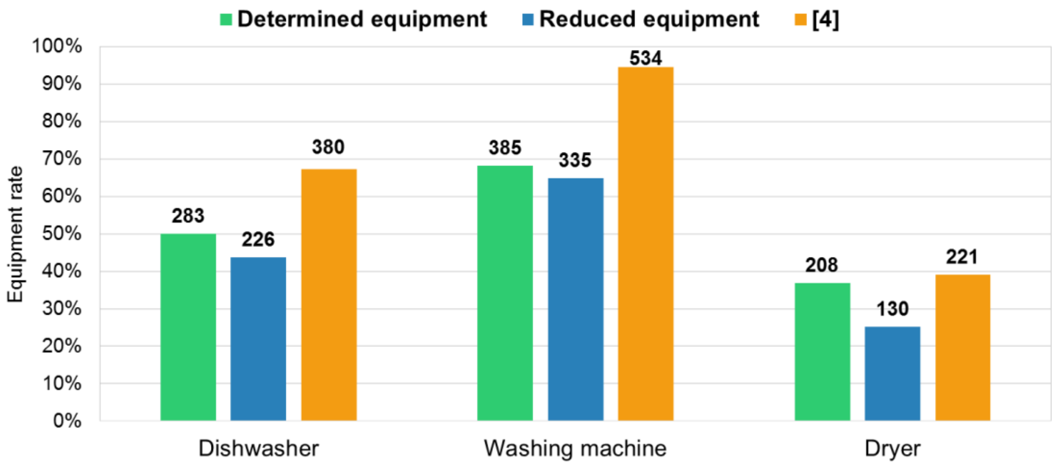

Figure 5: *Equipment rate of household appliances*

3.2 How often are the appliances used?

The number of uses per year has already been mentioned because it is used to modify the determined equipment rate. In Figure 6 the average amount of uses for every appliance is illustrated in the same structure as before (cf. Figure 5).

This analysis clearly shows the positive effect of the aforementioned restriction to a plausible range. For every appliance the reduced values are closer to the statistics. [5, 6] As before, the values differ slightly from the statistically determined ones, but are close enough to consider the results reliable.

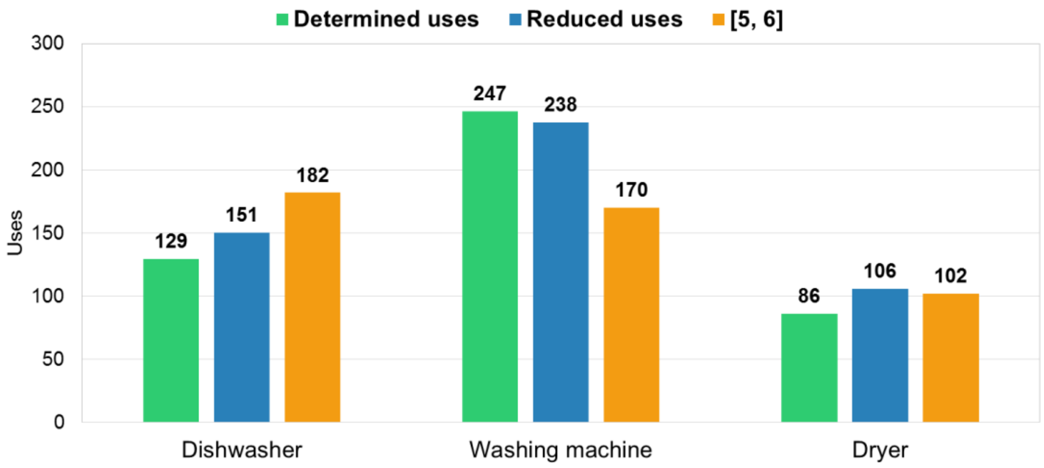

Figure 6: *Uses per year*

3.3 Which characteristics do the load profiles have?

A detailed investigation of all load profiles' characteristics cannot be given here, thus an exemplary load profile of an assigned dishwasher is used for demonstrating the characteristics as well as the functionality of the described algorithm.

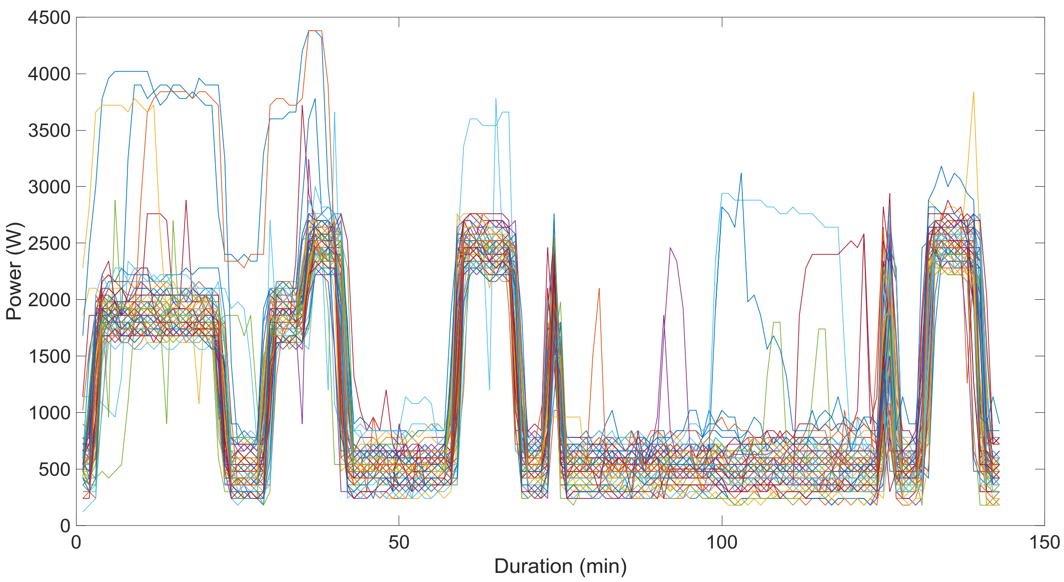

Figure 7: *Load profiles of dishwasher*

Figure 7 shows all matching patterns for a determined dishwasher profile in a whole year's load curve. The characteristics of this dishwasher can be clearly derived from the figure: all identified patterns consist six distinct heating phases of reproducible distance and length. Some of the patterns evince higher load, which stems from additional devices which are operating at the same time. Therefore, the presented method is not only capable of identifying isolated appliance uses. The red load profile shown in Figure 4 is the average of all identified load patterns contained in Figure 7.

3.4 What is the energy consumption per use?

The energy consumption per use is very important for determining the DSM potential. Moreover, these results can again be used for checking the plausibility of the applied method. The boxplot in Figure 8 shows the distribution of the appliances' energy consumption per use.

Previous studies mention the average energy consumption for dishwashers with 1.25 kWh, for washing machines with 0.75 kWh and for dryers with 1.6 kWh [7]. The comparison of these values and those depicted in Figure 8 (mean 1.01 kWh/0.42 kWh/1.58 kWh, median 0.90 kWh/0.43 kWh/ 1.38 kWh), leads to the result that the determined energy consumption is on average slightly smaller, but reasonably close. A possible explanation for the deviation might be that actual households tend to use eco programs, which is not considered in the computation of the average literature values.

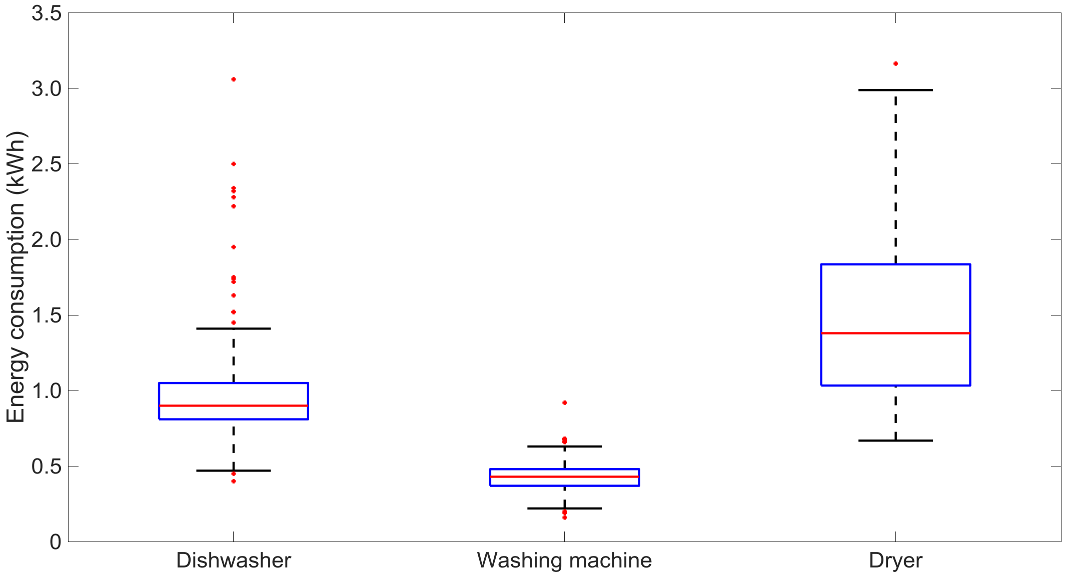

Figure 8: *Distribution of energy consumption per use*

3.5 How long does a program last?

The duration of the programs is the next value which belongs to the appliances' characteristics. Figure 9 shows the distribution of the appliances' program duration. According to the aforementioned source, program durations for dishwashers are in the range of 60 to 80 min, for washing machines between 70 and 120 min and for dryers between 80 and 120 min. The depicted boxplot confirms that most of the recognized dishwashers and dryers are in these ranges (mean 65 min/18 min/83 min, median 66 min/17 min/77 min). However, washing machines differ significantly. The analysis of the identified load profiles suggests that for washing machines, the described algorithm only recognized the initial heating phase, but not the subsequent spin cycle, since it is below the chosen power threshold. Since for the evaluation of DSM measures, the energy consumption is the important quantity, the results can nevertheless be used despite these shortcomings.

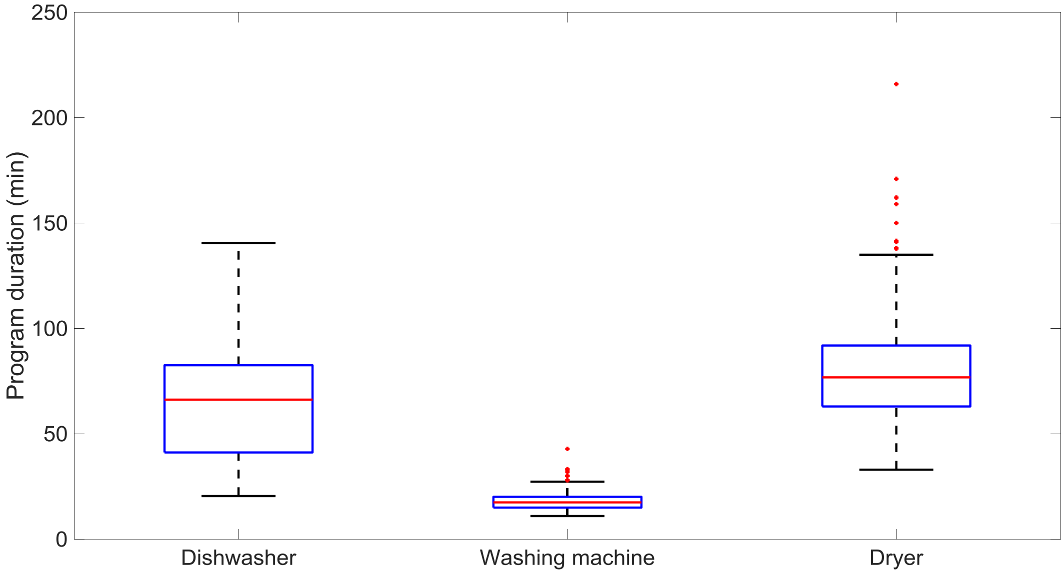

Figure 9: *Distribution of program duration per use*

3.6 When are the appliances used?

The last important subject of investigation is analyzing the user behavior, i. e. the typical times of operation of these appliances. These results are also essential for assessing the DSM potential. In this context, heat maps are used to visualize the distribution. The generated heat maps consist of 52 columns (one for each week of the year) and 168 rows (7 days with 24 hours each). Each cell contains the normalized number of identified operation of all 565 analyzed households during the relevant hour. This means that hours with no uses are marked black and the maximum is represented in white. Figure 10 shows the heat map for dishwashers for the year from Friday, August 1 2014 until Friday, July 31 2015.

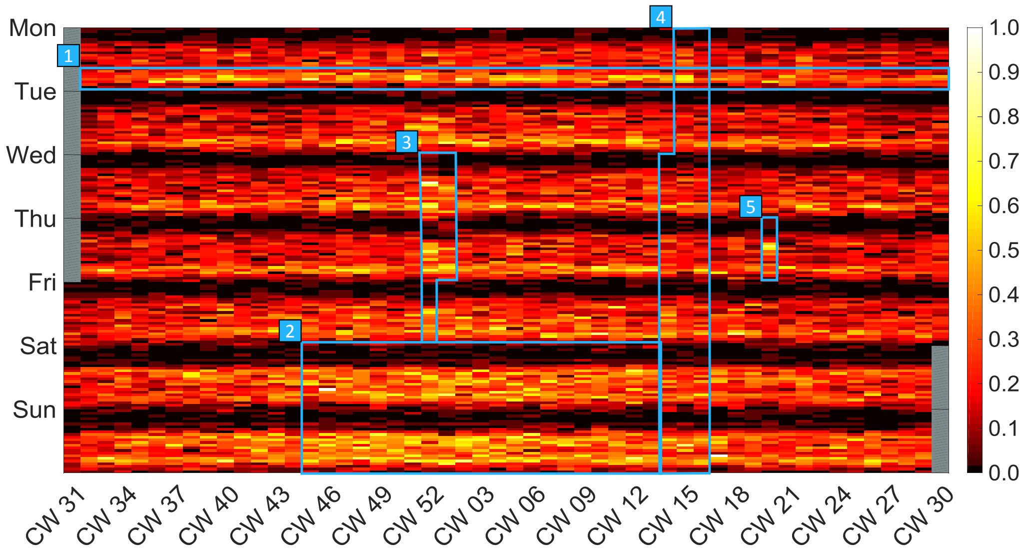

Figure 10: *Heat map – dishwasher uses*

This kind of visualization evinces some interesting characteristics. Five numbered blue boxes mark special periods of time:

1. Focus during working days: The first box marks the focus of uses during working days, in this example Mondays. It shows that most dishwashers are used in the evenings, which also holds

for other working days. Turning on the dishwasher after coming home from work is the expected user behavior for employed people. In addition, it has to be considered that about 14 % of the analyzed households has time-of-use tariffs, which incentivize usage in off-peak hours.

2. Focus during winter: The second area represents the focus of uses from November until March. The colored cells show a highly frequent use of dishwashers during weekends. A higher probability for being at home might be a conclusive explanation for this observation.
3. Christmas and New Year: The third box contains the following days: Christmas Eve, the 1st and 2nd Christmas day, New Year's Eve and New Year. The coloring clearly symbolizes increased uses of dishwashers. In comparison to the relevant days during the whole year these five days have a brighter color. It also makes sense that during these days, dishwashers are used more frequently.
4. Easter vacations: Easter holidays in the relevant federal state are highlighted with the fourth box. Compared to the rest of the year, the distribution of dishwasher uses is much smoother. This can be explained by the assumption that more people are on vacation, which reduces the after-work-peak in the evenings.
5. Christi Himmelfahrt (public holiday): The day emphasized with the fifth box is a public holiday in the considered federal state. This day differs from the other Thursdays. During this day there are more uses of dishwashers and these uses are also earlier compared to other working days.

All of these observations are in accordance with typically expected customer behavior. This is therefore another validation step of the presented algorithm.

In Figure 11 the heat maps for washing machines and dryers are depicted. As presented in chapters 3.1 and 3.2, both the equipment rate of washing machines and the number of uses per year are the highest. This can be seen in the intensive coloring of the relevant heat map. In contrast, the dryers' heat map is much darker which implicates the smaller equipment rate and the lowest number of uses per year.

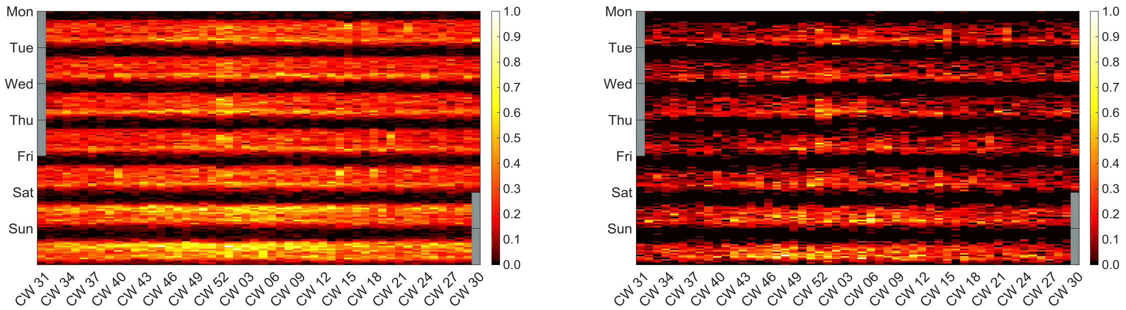

Figure 11: *Heat map – washing machine (left) and dryer (right) uses*

For both appliances, the focus of uses can be observed during weekends, for dryers especially in winter. This again seems plausible because people are doing their laundry rather on weekends. Moreover, higher dryer usage during winter makes sense because some households have other options to dry their laundry, e. g. a backyard.

4 Conclusion

The intention of the presented work is to develop an algorithm for disaggregating the load curve of a variety of households without having data about individual equipment rate or load profiles of the respective appliances. Load profiles are identified and extracted based on several criteria, such as power thresholds and durations. Measured load profiles for some appliances are used to derive these criteria, which allow assigning the identified load profiles to appliance types. Therefore, this method can be applied to a large number of data sets without the necessity to perform individual analyses or measurements.

The results show plausible and valid results. All investigated quantities, like equipment rates, energy consumption or number of uses are reasonable close to literature values. The investigation of the temporal user behavior even evinces noticeable differences for special days like holidays, which also confirms the validity of the method. The identification of dishwashers, washing machines and dryers can therefore be used for further work like the investigation of the households' DSM potential [8,9].

References

[1] Hinterstocker, Michael; Salvamoser, Simon; von Roon, Serafin; Berner, David; Staats, Joachim; Bruce-Boye, Cecil: Dynamische Korrektur der Lastprognose von Haushaltskunden mittels kurzzyklischer Smart-Meter-Daten. 9. Internationale Energiewirtschaftstagung, Wien, 2015

[2] Kelly, Daniel A.: Disaggregating Smart Meter Readings using Appliance Signatures. London, GB: Imperial College London, 2011

[3] Amon, Anatol: Analyse von Kundenlastprofilen und automatische Zerlegung in Gerätelastgänge. Wien: Technische Universität Wien, Institut für Energiesysteme und Elektrische Antriebe, 2011

[4] Ausstattung mit Gebrauchsgütern - Ausstattung privater Haushalte mit Haushalts- und sonstigen Geräten nach Gebietsständen im Zeitvergleich. In: https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/EinkommenKonsumLebensbedingungen/AusstattungGebrauchsguetern/Tabellen/A_EVS_Haushaltsgeraete_Gebietsstaende.html (aufgerufen am 20.06.2016) Archived by WebCite® at http://www.webcitation.org/6iPJtuer4. Wiesbaden: Statistisches Bundesamt (destatis), 2016

[5] Prior, Dirk: Nachbildung der Energiebedarfsstruktur der privaten Haushalte - Werkzeug zur Bewertung von Energiesparmaßnahmen in: Fortschr.-Ber. VDI Reihe 6 Nr. 379. Düsseldorf: Verband Deutscher Ingenieure (VDI), 1997

[6] Lünsdorf, Ontje: Selbstorganisation virtueller Geräte für das Lastmanagement von Kleinverbrauchern - Dissertation zur Erlangung des Grades eines Doktors der Naturwissenschaften. Oldenburg: Carl von Ossietzky Universität Oldenburg, 2012

[7] Dostal, Jakob: Statistische Auswertung sekundengenauer Lastprofile von Haushaltsgeräten. Wien: Institut für Energiesysteme und elektrische Antriebe; Technische Universität Wien, 2014

[8] Hinterstocker, Michael; Schott, Paul; von Roon, Serafin: Increasing residential self-consumption of PV energy by DSM. 10. Internationale Energiewirtschaftstagung, Wien, 2017.

[9] Hinterstocker, Michael; Schott, Paul; von Roon, Serafin: Evaluation of the effects of time-of-use pricing for private households based on measured load data. 14th International Conference on the European Energy Market, Dresden, 2017 (submitted).