Towards the Automated Extraction of Flexibilities from Electricity Time Series

Kaulakiene, Dalia; Siksnys, Laurynas; Pitarch, Yoann

Published in:
Proceedings of the Joint EDBT/ICDT 2013 Workshops

DOI (link to publication from Publisher):
10.1145/2457317.2457361

Publication date:
2013

Document Version
Early version, also known as pre-print

Towards the Automated Extraction of Flexibilities from Electricity Time Series

Dalia Kaulakienė
Department of Computer Science, Center for Data-intensive Systems, Aalborg University, Denmark
daliak@cs.aau.dk

Laurynas Šikšnys
Department of Computer Science, Center for Data-intensive Systems, Aalborg University, Denmark
siksnys@cs.aau.dk

Yoann Pitarch*
University of Lyon, CNRS, University of Lyon 1, LIRIS, UMR5205, F-69622, France
ypitarch@liris.cnrs.fr

ABSTRACT

Several recent and ongoing smart grid projects aim at incorporating more renewable energy sources (RES) into the energy production. Among them, the European MIRABEL project tackles this problem by managing flexibilities on energy demand and supply. Typically, this project assumes that some parts of the energy demand can be shifted when the RES production is sufficient, e.g., the washing machine can be turned on when the wind blows. To express these flexibilities, the project introduces the core-concept of flex-offer. Unfortunately, flex-offer data from the consumers is not yet available. Consequently, in order to test and evaluate the MIRABEL prototype, the flex-offers are extracted from the real world electricity consumption time series. In this work, we investigate, discuss, and experiment several ways to automatically capture flexibility within the electricity time series. Particularly, we show that incorporating domain knowledge, for instance, appliance information or appliance usage frequencies, can improve a lot the outcome of the flex-offer generation and, thus, the MIRABEL project global evaluation.

1. INTRODUCTION

There are many smart grid projects aiming to incorporate more renewable energy sources (RES) into the electrical grid. The problem of the RES production is that it does not follow the energy consumption. The RES production solely depends on the weather conditions, thus it can only be predicted, but not planned. In other words, the production from RES cannot be shifted to match the demand, e.g., the wind not necessary blows when you want to watch TV. On the other hand, some of the energy demand can be shifted in time to match the surplus RES production, e.g., the washing machine could be turned on when the wind blows.

*The work done while working at Aalborg University

EDBT/ICDT '13 March 18 - 22 2013, Genoa, Italy

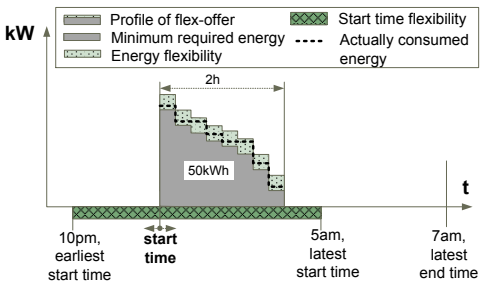

Figure 1: Example of a flex-offer

The described demand shifting approach is implemented in the smart grid project MIRABEL (abbr. Micro-Request-Based Aggregation, Forecasting and Scheduling of Energy Demand, Supply and Distribution) [1]. It assumes that energy consumer explicitly specifies his (or hers) flexible demand, e.g., "I want to wash my clothes my next morning". Moreover, the project explores such explicit flexibilities gathered from multiple consumers in order to schedule them in time so the flexible demand matches surplus production from RES [2]. Fig. 1 shows the example of the flex-offer issued by the owner of the electric vehicle. The flex-offer states that the charging of the vehicle's batteries should start between 10PM and 5AM, the charging takes 2 hours in total, and it requires 50kWh to be fully charged. The figure also shows the flexibilities in the energy profile, specifically, at each (15 min) time interval it states the minimum and maximum required energy (solid and dotted area respectively). The MIRABEL project prototypes the data management infrastructure which enables near real-time flex-offer collection [3], aggregation [4], and scheduling [5], in addition to the reliable and near real-time forecasting of energy production and consumption [6]. To sum up, the flex-offer concept is the basis of the project.

The consistent and reliable evaluation of the MIRABEL cannot be achieved without data from real world. Since real flex-offers from the consumers are not yet available (and their collection is out of scope of the project), the flex-offers are being randomly generated for the testing purposes. Specifically, the random approach assumes that consumption at every moment of a day is potentially flexible. Besides this assumption is very likely being false (e.g., flexi-

ble demand has few chance to exist during working hours), it can lead to dramatic consequences in the terms of the global evaluation of the MIRABEL system. For instance, with this random generation strategy, we can hardly analyze the scalability of MIRABEL during the peak hours since macro (or aggregated) flex-offers are more or less uniformly dispatched within the day. To conclude, the project needs a tool which generates realistic flex-offers complying with the realistic assumptions of consumers' and producers' behavior in the electricity market. The tool could incorporate the domain knowledge in either of two ways: (1) to detect candidate periods for flexible demand and (2) to estimate the amount of flexible energy. Generally, the electricity consumption time series exhibit 0.1–6.5% of flexible demand [7]. On the other hand, the production time series flexiblities highly depend on the concrete producer and its business model. For this reason, in this paper we focus solely on the flexibility extraction from the consumption time series. To sum up, in this paper we (1) propose the design for flexibility extraction tools at appliance level and the total household consumption level, (2) evaluate the proposed extraction approaches in terms of how realistic the extracted flex-offers could be, and (3) specify directions for the future research.

The paper is organized as follows. The general flexibility extraction process is provided in Section 2, then subsequent sections describe two types of extraction approaches – the first ones operate on the total household consumption level to produce household flex-offers (Section 3) and the others aim to decompose the total consumption into appliance level consumption and produce appliance level flex-offers (Section 4). Lastly, the related work is presented in Section 5 and the conclusions and future work are given in Section 6.

2. GENERAL FLEXIBILITY EXTRACTION APPROACH

Given historical consumption time series, the objective is to extract the flexible energy from it and capture the extracted flexible energy in respective flex-offers. The potential flexibilities in the historical time series highly depend on the individual energy consumer. Hence, a lot of domain knowledge about time series need to be taken into account in order to extract realistic flex-offer from that.

The general architecture of the flex-offer extraction is presented in Fig. 2. The flex-offer extraction algorithm relies on some context assumptions, and, based on these assumptions, it decomposes the given historical time series into flexible and non-flexible parts. The input of flexibility extraction is historical time series and the context information, including specific parameters for the specific extraction algorithm. In the extraction process, the input time series are analyzed, taking into account the contextual assumptions and the context information, and then the potential flexibilities are extracted, formulated as flex-offers and outputted together with the modified time series (the flexible energy extracted from the original ones).

We propose two types of flexibility extraction approaches (others might exist) visualized in Figure 3. The first type operates at the total household consumption time series and it outputs flex-offers which represent the total flexibility potential (left branches of the hierarchy). The second type aims to decompose the input time series into appliance level consumption and then outputs appliance level flex-offer which represent potentially flexible usage of one appliance at the household (right branches of the hierarchy). The next sections will elaborate the design of the five approaches of the household flexibility extraction.

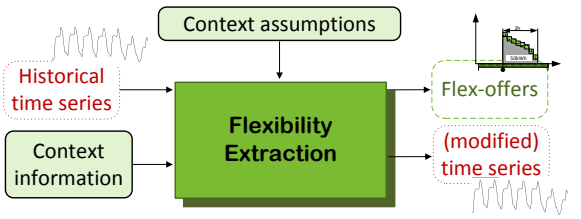

Figure 2: General approach of a flex-offer generator

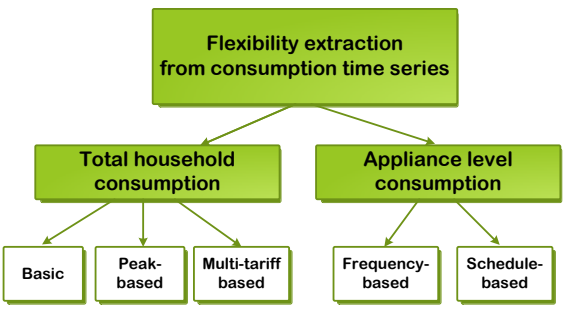

Figure 3: Approaches of flexibility extraction from household consumption time series

3. FLEXIBILITY EXTRACTION FROM THE TOTAL AGGREGATED HOUSEHOLD CONSUMPTION TIME SERIES

The extraction of this class operates on the aggregated energy level, i.e., both the input time series and the extracted flex-offers represent total aggregated consumption of many smaller appliances in the household.

3.1 Basic Approach

The basic approach assumes that some percentage of the household consumption is flexible and this flexibility is available at any time of the day. However, although this assumption is not very realistic, as discussed in Section 1, it can state the basis of flexibility extraction and flex-offer formulation process. The inputs and context assumptions to the basic extraction process is summarized as follows.

Context assumptions: At any given time of the day, some of the household consumption is flexible.

Input time series: The basic extraction accepts historical time series, representing total household consumption composed of many appliances at the household.

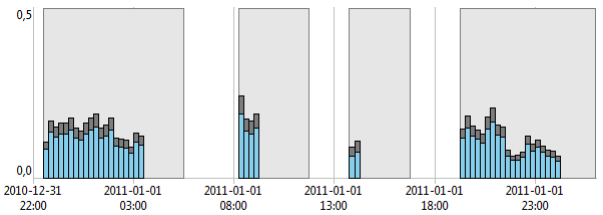

Figure 4: Flex-offers extracted using the basic extraction approach

Context information: The basic extraction expects some parameters. The most important is the percentage of the flexible demand part in the input time series. Other parameters are directly related to the flex-offer attribute information (see the description of Figure 1), e.g., the number of intervals in a single flex-offer, interval duration, minimum and maximum percentage of required energy, creation time, acceptance time, assignment time, earliest start time, and latest start time. All these parameters are randomized in controlled variation limits in order to generate non-uniform flex-offers.

The process of the flexibility extraction starts with the division of input time series into periods, and then one flex-offer is extracted for each of the periods spanning few hours, then the fraction of flexibility within each period is calculated (based on the configuration parameter). Lastly, a flex-offer for each period is extracted. Afterwards, time and energy amount flexibilities are built by applying some randomization to the constructed flex-offers.

We have implemented the basic flex-offer extraction approach (see the resulting flex-offers in Figure 4). Light areas represent minimum required energy, and the dark shaded ones represent maximum required energy. In the figure, there are 4 flex-offers, occupying their own periods of the time axis. For each flex-offer, the total energy amount (the sum of the average required energy in the profile intervals) is equal to the flexible part extracted from the input time series. Although the resulting flex-offers have all the required attributes and all of these attributes are within the required limits, the actual quality of the output is not known. There exist no real flex-offers in the world, thus, the statistics (e.g., correlation, sparseness, autocorrelation) of the output of flexibility extraction cannot be evaluated. In addition to this basic idea, more sophisticated algorithms could be applied, for instance, to analyze which part of the daily consumption is flexible, or which part of the taken period is the actual flex-offer and which is the time shifting flexibility.

3.2 Peak-based Flexibility Extraction

This approach takes into account the intuition that, during the energy consumption peak, there are more appliances using the electricity, thus, there is larger probability that at least one of these devices exhibits flexible usage. Also, the peak-based approach assumes that a consumer uses one flexible appliance per day, thus, only one flex-offer per consumer per day should be extracted. Moreover, the flex-offer should be positioned on the time axis when the largest consumption

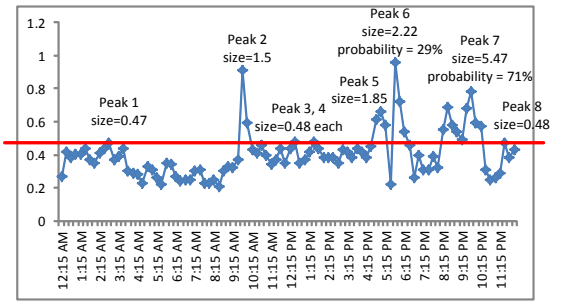

Figure 5: Example of the peak-based extraction approach. Firstly, candidate peaks are detected, then filtered, and finally, the single one is randomly selected

peak is detected, as suggested by name.

Context assumptions: During the energy consumption peaks, there is more room for flexibility since more appliances (potentially) contribute to the total energy consumption. Another simplifying assumption is the existence and the extraction of one flex-offer per consumer per day.

Input time series: The peak-based extraction accepts historical time series, representing total aggregated household consumption.

Context information: The peak-based extraction expects the same parameters as the basic extraction approach, i.e., flexible demand part during the day and specific parameters for formulating resulting flex-offers.

Figure 5 shows the consumption time series from one household during one day. The peak-based approach starts by detecting peaks in the 24-hour period of the household consumption. The peak detection process firstly calculates the average daily consumption and considers only those peaks which have energy amount greater than average during the whole period. It is shown in Figure 5 as a thick horizontal line at around 0.46 kWh. Then the peak filtering phase discards some peaks, which have the total energy amount smaller then the flexible part of the day. For example, if the flexible part (expressed as the context parameter) is 5%, then the flexible part of the energy of the day shown in the figure is 39.02 * 0.05 = 1.951 kWh. Therefore, the peaks 1, 2, 3, 4, 5, and 8 have to be discarded, because their total energy amounts (denoted as peak size in the figure) is smaller than 1.951. The next step in the peak-based approach is the selection of the single peak where the output flex-offer will be positioned in the time axis. The remaining candidate peaks after the peak filtering (peak 6 and 7 in the figure) are given probabilities of being selected depending on their size (peak 6 – 29%, peak 7 – 71%), and the single peak is randomly chosen depending on these probabilities. Finally, the flex-offer is generated using the same methodology as in the basic approach.

The peak-based approach is more realistic than the basic one, but the assumption about one flex-offer per consumer per day might not be true for all the consumers — some would possibly have more flexible consumption (many flexible usage appliances) and the others may have less (e.g., only one washing machine for 2 persons household).

3.3 Multi-tariff Based Flexibility Extraction

The multi-tariff approach explores the fact that consumers change their electricity consumption behavior when the multi-tariff (also called variable rate) billing system is introduced. In other words, consumers delay the flexible usage (e.g., washing machine) to the low tariff time (e.g., after 10PM). Later we refer to the multi-tariff billing system as multi-tariff period. The input and output of the multi-tariff extraction process is summarized as follows.

Context assumptions: Consumers change their behavior when the multi-tariff system is introduced. Specifically, they delay their flexible consumption to the low tariff time.

Input time series: The input for the multi-tariff extraction should consist of two historical time series – first time series should contain the consumption of the consumer during one tariff period, and the second one should contain consumption from the same consumer during the multi-tariff period. Note that flexibility is extracted only from the multi-tariff time series, and the one tariff time series are used only for reference.

Context information: The multi-tariff extraction only expects parameters needed for the generation of output flex-offers.

Output: Since the multi-tariff extraction accepts two types of time series, its output also differs from the previous approaches. It outputs unchanged historical time series representing consumption during one tariff billing system. It also outputs flex-offers extracted from the multi-tariff time series and the modified multi-tariff time series where the flexible energy amount is subtracted.

The multi-tariff approach firstly analyzes one tariff time series to estimate the usual consumption of a consumer. It can calculate the typical behavior during the work days, weekends, holidays, different seasons of the year, etc. Then, the extraction approach takes multi-tariff time series and detects the flexible consumption in it by comparing with the typical consumption in one tariff. The multi-tariff approach has very realistic assumptions and, probably, extracts very realistic flex-offers. Unfortunately, we do not have the required time series for this approach, thus, we cannot show any results of it.

4. FLEXIBILITY EXTRACTION USING INFORMATION ABOUT APPLIANCES

An appliance level flexibility extraction is designed to extract flex-offers which represent the potential usage of one single appliance. The general architecture of this approach

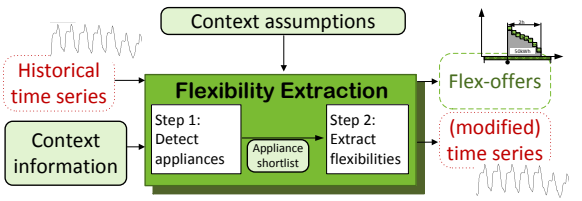

Figure 6: General approach of the appliance based flexibility extraction

Appliance name	Energy Consumption Range (KWh)	Energy profile
Vacuum Cleaning Robot from Manufacturer X	0.5 - 1	...
Washing Machine from Manufacturer Y	1.2 - 3	...
Dishwasher from Manufacturer Z	1.2 - 2	...
Small Electric Vehicle	30 - 50	...
Medium El. Vehicle	50 - 60	...
Large El. Vehicle	60-70	...

Table 1: Example of an appliance information

is shown in Figure 6. The only difference from the general architecture is the flexibility extraction process – it is composed of two steps. Firstly, the appliances, contributing to the total consumption, are detected, and then this information is used in the second step, where the flexibilities from the input time series are actually extracted. Also, the appliance level extraction relies on the specifications of the electricity consumption of all possible appliances in fine-grained manner. An example of such information is given in the Table 1. Here, for each manufactured appliance, we define their energy consumption ranges and energy profiles with min and max ranges for every time stamp (granularity must be even smaller than 15min).

4.1 Frequency-based Appliance Level Flexibility Extraction

The frequency-based flexibility extraction is based on the observation that some of the appliances may be used daily while some may be used weekly or monthly, or even yearly. The input of the frequency-based extraction process is summarized as follows.

Context assumptions: The consumption time series is composed of the consumption of many appliances. Given the known consumption profiles of appliances (and perhaps other specification) from manufacturers, it is possible to derive the set of the appliances which contributed to the provided time series.

Input time series: The input is total historical household consumption time series.

Context information: The specification of the electricity usage of all appliances ever manufactured in the world.

In addition, the frequency-based extraction approach expects parameters needed for generation of output flex-offers, i.e., creation time, acceptance time, assignment time, earliest and latest start time.

Output from the step 1: The step 1 of the extraction derives the shortlist of the possibly used appliances and their frequency usage table.

Output from step 2: The step 2 outputs a set of extracted flex-offers, each of them corresponding to one usage of a specific appliance at a specific time period. The extraction also outputs modified historical time series where the flexible energy amount (contributing to the energy amount in flex-offers) is subtracted.

The frequency-based extraction starts with the step 1. It takes two types of input data, the specifications of electricity consumption of many possible appliances and the historical consumption time series at the household level. Then it applies various data mining and machine learning algorithms to derive which appliance and how frequently was used. The output of the step 1 is a shortlist of the possibly used appliances, their usage frequency, and the time flexibility (difference between latest start time and earliest start time). Then step 2 takes the original historical time series and the shortlist, and it distributes possible "activations" of the appliances respecting the usage frequencies. For instance, a vacuum cleaning robot (like iRobot Roomba) cleans the house every day at 10AM, and afterwards, but before the next day, it needs to charge its batteries. Thus, the shortlist containing vacuum cleaning robot provides its frequency as once per day, and time flexibility as 22 hours (it needs to be charged before the next usage).

This approach extracts very realistic flex-offers, since it decomposes the total consumption time series into the consumption of individual appliances.However, the decomposition is a non-trivial task. Most of the research is targeted towards the prediction of the future usage [8], but similar methods could be applied to derive the historical usage pattern [9]. Moreover, the output of the step 1 of the extraction can be reused for other households which exhibit similar consumption characteristics.

4.2 Schedule-based Flexibility Extraction

The frequency-based generator can be further extended to take into account the habits of each family. The insight is that the information about the frequency of the appliance usage is very simplified, for instance, the dishwasher is more used during the weekends since the family eats at home more often than during the workdays. The input and output of the extraction is summarized as follows.

Context assumptions: The consumption time series is composed of the aggregated consumption of many appliances. Given the known consumption profiles (and perhaps other specification) from the manufacturers, it is possible to derive the set of the appliances which contributed to the provided time series. Moreover, the usage of the appliances is not uniform, thus, the exact schedule of the usage of each appliance can be derived.

Input time series: The input is the time series of the total historical household consumption.

Context information: The context information is the same as for the frequency-based extraction. Firstly, the specification of the electricity consumption of all appliances the household could potentially use. In addition, the parameters for generation of output flex-offers.

Output from the step 1: The step 1 of the extraction derives the shortlist of the possibly used appliances and their usage schedule.

Output from step 2: The step 2 outputs the list of extracted flex-offers, each of them corresponding to one usage of a specific appliance at a specific time period. The extraction also outputs modified historical time series where the flexible energy amount (contributing to the energy amount in flex-offers) is subtracted.

The approach of the schedule-based extraction is similar to the frequency-based one. Firstly, it derives the shortlist of the appliances and their usage schedule. Then in step 2, the extraction formulates flex-offers based on the given schedule and other specific parameters. However, in order to derive the schedule of the appliances, the total consumption time series must be decomposed into the consumption of each contributing appliance. All the problems, stated for the frequency-based extraction still holds. Moreover, the schedule detection is a more complex task than the frequency detection [10]. On the other hand, the output flex-offers are very realistic since they are based on the deep domain knowledge about factual energy consumption profiles of concrete appliances and their potential usage schedule.

5. RELATED WORK

The problem, presented in this paper, is related to the general problem of the synthetic data generation. There are different approaches depending on problem domain and the data type. Generally speaking, synthetic data generation problem is so common that even general analytical tools, such as Matlab [11], include capabilities of generating data depending on the given configuration parameters.

The time series generation for the energy domain is an emerging field. However, most of the research is done in synthetic RES production data generation, since they are beneficial for the classical energy market and power systems. Although the consumption time series usually exhibit some patterns (especially, at aggregated level), there is a lack of consumption time series generators which incorporate some domain knowledge, e.g., the typical electricity consumption of the two resident household or a family living in a suburb.

The other problem, analyzed in this paper, is related to the pattern extraction from the given time series. Usually, the time series is composed of the trend, seasonal, and error components [12]. Also, there exist more sophisticated time series data mining techniques, which stress subsequent matching, anomaly detection, specific feature extraction [13]. Moreover, time series disaggregation algorithms are applied for reasoning about the finer granularity of the

data than the input, e.g., filling the missing values [14]. However, data mining techniques for flexibility extraction (or detection) is orthogonal task. Data mining techniques could help to recognize the time periods were flexible consumption was possible, i.e., the usage of a particular appliance [9], but these algorithms are unable to estimate the minimum-maximum bounds of required energy, or time flexibilities of the potential flex-offer.

6. CONCLUSIONS AND FUTURE WORK

This paper presented the problem of extracting flexibilities from historical energy consumption time series. We discussed the feasibility of the proposed 5 approaches for extracting flexibilities from the household consumption time series. The approaches operating on the total household consumption level rely on less realistic assumptions, such as some percentage of the energy consumption is flexible during any period of the day. The appliance level extraction approaches, on the other hand, rely on the more realistic assumption that the total household consumption is composed of the consumption from many appliances at home. These extraction algorithms aim to employ data mining and machine intelligence techniques to derive the appliance activation periods and, based on that, extract the potential flexibilities. Appliance level extraction approaches are more sophisticated and very realistic, but the actual implementation of them is left as a future work since the granularity of the available time series is not sufficient (only 15min granularity). Furthermore, the appliance level extraction approaches can be easily extended to the real-time flex-offer generators, which detect flexibilities and formulate flex-offers based on the usual appliance usage or the given (mined) schedule of the household.

Regarding the MIRABEL scenario, individual flex-offers have to be aggregated from thousands consumers before the actual scheduling (and matching with the surplus RES production). The aggregated flex-offer should exhibit similar patters as the ones extracted directly from the aggregated time series from thousands consumers. Despite the fact that the peak-based approach produces not very realistic flex-offers, the aggregated flex-offers are pretty realistic. Currently, the peak-based approach is used for the MIRABEL evaluation.

Further research directions include flexibility extraction from industrial consumers. Since the flex-offer concept is designed to capture flexibilities in both production and consumption, the next step is to design and implement production flexibility extraction algorithms. It should be noted, that production flexibilities exhibit somehow different assumptions than consumption flexibilities. As mentioned earlier, RES production, usually, is not flexible in a sense that it highly depends on the weather conditions. On the other hand, RES producer can maintain highly specialized and accurate local weather forecast and foreseen that wind will be sufficiently strong in two hours ahead. As a result of this forecast, the RES producer could issue a production flex-offer specifying that the start of electricity production can be either in 2 hours or 3 hours ahead, depending on the flex-offer schedule. Traditional electricity producers are even more flexible, thus, they can issue production flex-offers for almost all of their production. In other words, the flex-offer concept enables the change in the current electricity market and shift current trading model based on bids to the explicit flexibility trading model. To sum up, the future research should be focused on covering the full spectra of energy producers and consumers within the electricity grid and exploring their flexibilities, i.e., generating flex-offers on the fly.

Acknowledgment

The work presented in this paper has been carried out in the project MIRABEL funded by the European Commission under grant agreement number 248195.

7. REFERENCES

[1] *The MIRABEL Project*, 2012. `http://www.mirabel-project.eu`.

[2] M. Boehm, L. Dannecker, A. Doms, E. Dovgan, B. Filipic, U. Fischer, W. Lehner, T. B. Pedersen, Y. Pitarch, L. Siksnys, and T. Tusar, "Data management in the MIRABEL smart grid system," in *Proc. of EnDM*, 2012.

[3] L. Šikšnys, C. Thomsen, and T. B. Pedersen, "MIRABEL DW: Managing complex energy data in a smart grid," in *Proc. of DaWaK*, 2012.

[4] L. Šikšnys, M. E. Khalefa, and T. B. Pedersen, "Aggregating and disaggregating flexibility objects," in *Proc. of SSDBM*, 2012.

[5] T. Tušar, L. Šikšnys, T. B. Pedersen, E. Dovgan, and B. Filipič, "Using aggregation to improve the scheduling of flexible energy offers," in *Proc. of BIOMA*, 2012.

[6] U. Fischer, D. Kaulakienė, M. Khalefa, W. Lehner, T. B. Pedersen, L. Šikšnys, and C. Thomsen, "Real-time Business Intelligence in the MIRABEL Smart Grid System," in *Proc. of BIRTE*, 2012.

[7] G. Černe, Z. Marinšek, H. Berthold, A. Savinov, C. Nychtis, H. Frey, L. Šikšnys, T. Tušar, and M. Boehm, "D6.2 Specification and Design of trial cases," 2011. `http://wwwdb.inf.tu-dresden.de/miracle/D6.2SpecificationandDesignoftrialcases.pdf`.

[8] K. Basu, M. Guillame-Bert, H. Joumaa, S. Ploix, and J. Crowley, "Predicting home service demands from appliance usage data," in *Proc. of ICTA*, 2011.

[9] U. Greveler, B. Justus, and D. Loehr, "Forensic content detection through power consumption," in *ICC*, 2012.

[10] T. Bapat, N. Sengupta, S. Ghai, V. Arya, Y. Shrinivasan, and D. Seetharam, "User-sensitive scheduling of home appliances," in *Proc. of Green networking*, 2011.

[11] Mathworks, "Matlab." `http://www.mathworks.se/products/matlab/`, 2012.

[12] M. E. Khalefa, U. Fischer, and T. B. Pedersen, "Model-based integration of past & future in TimeTravel," *PVLDB*, vol. 5, no. 12, 2012.

[13] J. Lin, E. Keogh, S. Lonardi, and P. Patel, "Finding motifs in time series." `http://cs.gmu.edu/~jessica/Lin_motif.pdf`, 2002.

[14] S. Vitullo, "Disaggregating time series data for energy consumption by aggregate and individual customer." `http://epublications.marquette.edu/dissertations_mu/169`, 2011.