Mining Flexibility Patterns in Energy Time Series from Industrial Processes

Nicole Ludwig, Simon Waczowicz, Ralf Mikut, Veit Hagenmeyer

Institute for Applied Computer Science, Karlsruhe Institute of Technology
E-Mail: nicole.ludwig@kit.edu

1 Introduction

The transition from traditional energy sources to renewable ones is not trivial. Various components in today's energy system are built for traditional suppliers and cannot cope with the new requirements imposed by the use of renewable energy sources (RES). The most commonly mentioned aspect differentiating renewable from traditional energy sources is the intermittent power generation. Although enough power can be supplied from renewable energy sources on average, the power is not always generated when actually demanded. While changing the supply strategy is not an option as RES depend on e. g. wind and sunshine, changing the demand side is possible. Methods to change the energy demand behaviour are usually summarized under the term *Demand Side Management* (DSM). For these management strategies to work sufficiently, the consumer's flexibility has to be determined, as there is the possibility for potential usage shifts. DSM potential has been analysed in great detail for households as well as some energy-intensive industry processes, however industrial batch processes have not yet been considered. As a first step in detecting demand side flexibility potential, this paper investigates the mining of patterns in energy time series data from industrial processes and introduces a new way of properly finding reoccurring motifs in industrial energy data.

Motif discovery, for example the Mueen-Keogh algorithm [9], has been applied to a great variety of problems, e. g. seismology [2], outlier detection [7], activity detection from sensor data [1], classifying heart sounds [6] and audio-based music structures [12]. In the realm of energy systems, [13] use motif discovery to reduce the dimensionality of large time series data sets and reduce the prediction errors when forecasting energy consumption. More prominently, motif discovery is used to identify individual appliances in energy consumption data. Some examples are [4] who are able to disaggregate the energy consumption in an household from a single-point of entry, but need a training phase of one week to find the characteristics in the household. [5] give an overview over more methods for disaggregation of an end-users smart meter data. However, all the methods mentioned work supervised and thus need labelling of the data.

Motif Discovery algorithms are very efficient at finding similar patterns in time series data. They excel if the found motifs are of approximately the same length. In energy time series this is the case for e. g. daily load curves from residential buildings. However, given industrial production data, the algorithms find proper motifs, but cannot identify the correct starting points if the motifs vary greatly in their length.

This paper uses a novel two-stage algorithm to correctly identify reoccurring motifs in industrial process time series data. We first use an event search algorithm to find possible starting points for the motifs and then use a standard motif discovery algorithm to identify which of those events are triggered by the same process. It is the aim to evaluate those motifs e. g. process lengths, energy intake etc. and classify the processes according to their degree of potential flexibility. The variability in the motifs is used as an indicator for potential flexibility but does not guarantee that this flexibility can also be used.

The remainder of the paper is structured as follows. Section 2 introduces the used motif discovery and event search algorithms, before Section 3 describes the found motifs and categorizes the resulting motifs in terms of their flexibility. Afterwards we discuss the implications (Section 4) before concluding in Section 5.

2 Methodology

The approach we chose can be split in four parts; first we discover where the *events* start with the help of a minimum search algorithm, then we find how a usual motif looks like with the help of a univariate motif discovery, afterwards we compare the motif found to the processes after the events, establishing whether they are similar. And lastly, we

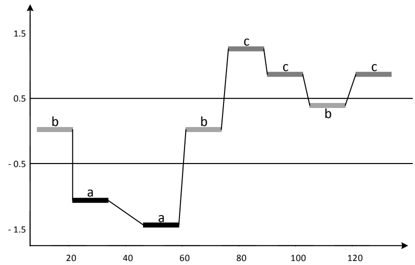


Figure 1: SAX transformation of an example time series.

categorize the differences in the found motif according to their flexibility potential.

In this section we first introduce a common motif discovery algorithm before explaining our event search approach. The notion of flexibility as well as the measures we use to describe it will be introduced thereafter.

2.1 Univariate Motif Discovery

Patterns in the load profile of the different buildings can indicate a potential to shift electricity demand or find times in which we could reduce the demand through energy efficiency measures. Especially deviations from a specific pattern but also variations in the same pattern can be good indicators for these demand side management or flexibility potentials.

Thus, we want to model normal behaviour for the steam demand of every building to be able to later detect deviations from this normal shape. While traditional clustering approaches tend to be slow on larger data sets, [3] developed an algorithm which can detect reoccurring patterns in time series while scaling very well and being robust to noise. This so-called *motif discovery* can be applied to many applications, but has not yet been applied to steam consumption load patterns for the purpose of finding demand side management potentials.

The algorithm is based on the *symbolic aggregated approximation* (SAX) of time series. For this approximation, the normalised time series are discretised and transformed into equal length *words* which are part of an *alphabet* with predefined length. With the help of a *sliding window* a matrix $S^* \in (n - m + 1) * w$ is generated, with n being the number of observations, m the number of subsequences and w the word length. The SAX representation of all subsequences, i. e. the words, are saved row-wise in this matrix. In every iteration of the algorithm, we randomly select l of the w columns of S^*, where l is a user-defined mask length and $l \leq w$. The word built with l columns is compared to all $(n-m+1)$ rows of S^*. If there exists similarity, the corresponding entry in the collision matrix is incremented. The entries with the highest values in the collision matrix are considered potential motifs. Those motifs are then iterated over the original time series and their distance is calculated to find the instances where the motif occurs. The distance measure here is a simple euclidean distance, but could also be e. g. dynamic time warping.

2.2 Event Search

Motif discovery often fails if the motif is not behaving regularly, particularly variation in the length of the process are hard to detect for fixed word or window lengths. While Industrial processes, especially chemical ones are often repeated batch processes, the batch processes do not have to be exactly the same for every iteration, but can vary in length and intensity. As we struggle finding proper starting points with the fixed window approach of the motif discovery described before, we want to find the starting points separately. We chose to do a simple minima search in a predefined window, to find the supposed starting points for each batch process and save the time series subsequence between two minima as our batch process. We treat the subsequences found equally to the motives found with the motif discovery algorithm before.

2.3 Dynamic Time Warping

ToDo

Dynamic time warping (dtw) is a method to find the optimal alignment given two time-dependent sequences.

For further explanation see e. g. [10]

2.4 Flexibility

Given we found proper motifs with the above described methods, we want to examine their potential flexibility. We work with the flexibility definition of [11]. They describe flexibility as *"the amount of energy and the duration of time to which the device energy profile (energy flexibility) and/or activation time (time flexibility) can be changed."* Thus, translating this to a data-driven approach; there are several cases, where we identify a potential degree of flexibility.

1. **Energy Flexibility**. The pattern always starts to occur at the same time and weekday but the length and intensity varies. This would indicate that the process can run in different modi and there is thus some degree of flexibility in the decision on the modi.
2. **Time Flexibility**. The same pattern in the time series occurs in the same form at different e. g. times of day or days of week, depending on the length of the pattern. For example, a process always runs Mondays, but the time varies greatly, this would indicate some level of flexibility regarding the starting time of the process on Mondays.
3. A combination of the above mentioned cases is also possible.

We describe the measures we use to judge those types of flexibility in the following.

2.4.1 Energy Flexibility

As stated before, the energy flexibility is mainly concerned with variations in the motif itself as opposed to when the motif occurs. We have a look at three different measures: the length, intensity and ramping time of the motifs.

The length is described by the time steps between the current local minimum and the next local minimum found through the event search.

$$\text{length} = t(\text{loacalmin}[j]) - t(\text{localmin}[i]) \quad \text{for} \quad i < j \tag{1}$$

The intensity is measured as the area under the curve, thus the energy used by the process. We approximate the area under the curve (AUC) using the composite trapezoidal rule. Given an interval $[a, b]$, we split this interval in M subintervals $[x_k, x_{k+1}]$ of equal width $h = (b-a)/M$ by using the space nodes $x_k = a + kh \quad \forall k = 0, 1, \ldots, M$. The composite trapezoidal rule for M subintervals can then be expresses as

$$T(f, h) = \frac{h}{2} \sum_{k=1}^{M} (f(x_{k-1}) + f(x_k)), \tag{2}$$

which is an approximation for the integral of $f(x)$ over $[a, b]$

$$\int_a^b f(x)dx \approx T(f, h). \tag{3}$$

We compare the ramping of the motifs by measuring the time steps needed between the local minimum, thus starting point, and the local maximum of the time series.

$$\text{Ramping} = t(\text{localmax}) - t(\text{localmin}). \tag{4}$$

2.4.2 Time Flexibility

In contrast to the energy flexibility described before, the time flexibility is only concerned with the starting times, days of week and length, thus end times and days of week. We assume that if a process runs every Monday at 8am it has to run at this specific point in time, while running always at a different day of the week would indicate that we have a certain flexibility here, as to when the process has to start. This is obviously a simplification as we neglect dependencies from other processes at the moment.

2.5 Framework

We use three steps to get to a categorisation of flexibility from steam consumption data. The whole framework is graphically shown in Figure 2. We start with the event search algorithm, saving the subsequences between each event as possible motifs. Afterwards, we run the motif discovery algorithm using the average length of the event subsequences as the word

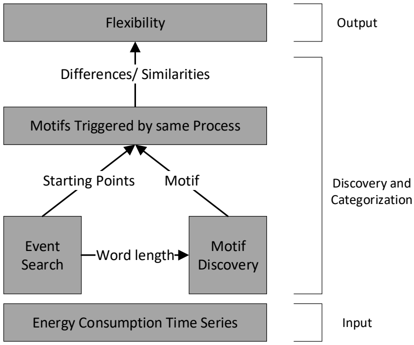


Figure 2: The framework used in this paper to extract flexibility potential from steam consumption data via motif discovery.

and window length. We then measure the similarity of the found motifs with the subsequences, using dynamic time warping. The motifs which are similar enough are used to examine their differences more closely and categorize their flexibility potential.

3 Results

To evaluate the proposed method we have a look at the steam consumption time series from a building of a chemical factory. The steam consumption is measured as hourly average in tons over a period of five weeks. Due to commercial sensitivity issues, no true measures or time indicators will be presented in this paper. Although we do not have information on the products manufactured, we know that there are possibly one or several repeating processes in the factory building. Given this consumption data we want to find the shapes of those recurring processes. The time series is depicted in Figure 3. The graph indicates a pattern which repeats itself on different levels of steam intake.

This section presents the individual steps described in the framework, starting with the motif discovery and event search, before analysing and categorizing the flexibility.

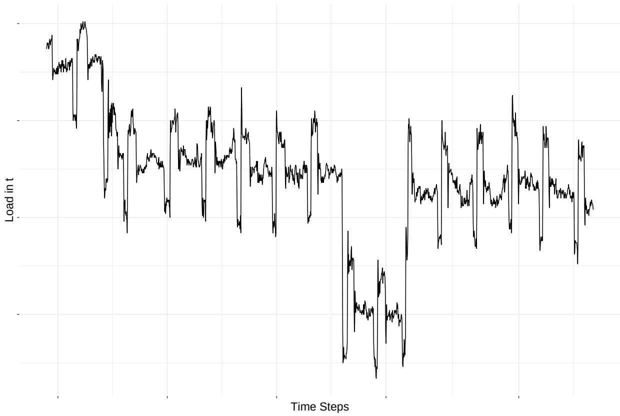


Figure 3: Original time series data of a buildings steam demand over five weeks.

3.1 Motif Discovery

Following the framework we have established, we run the above described minima search algorithm with a window size of 100 and store the local minima found in a vector. We then build subsequences of the time series, starting at each of those minima.

The minima search gives a very good result to find the different time series subsequences which look similar enough to be from the same batch process. Figure 4 shows some exemplary subsequences found through the event search. The structure of the process seems to remain the same over the different instances. However, the length and intensity of the process varies, which is useful for our purpose as it could be an indicator for flexibility. Figure 5 shows a heatmap of all the subsequences, where we have normalized all subsequences with mean $\mu = 0$ and standard deviation $\sigma = 1$. The subsequences are aligned at their minimum point. We can clearly see that they exhibit a similar structure but differ in length.

Next, we run the motif discovery algorithm using the additional information we have about the processes. We chose a relatively small alphabet size of $a = 10$ as the variations in patterns do not seem to be high. Choosing the word length is more difficult as it has a great impact on the found motifs. Additional information on the *normal* length of the

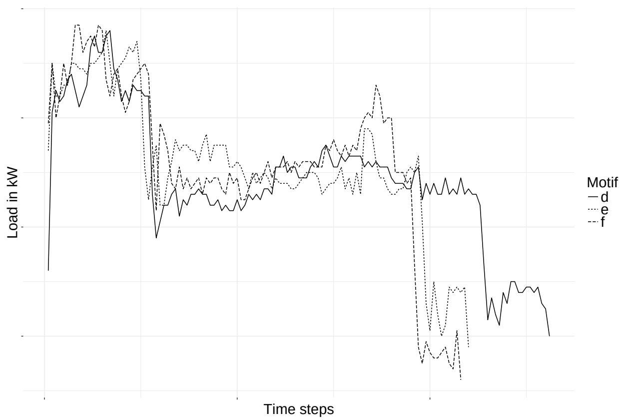


Figure 4: Exemplary subsequences of the steam consumption time series. Each subsequence starts after a minimum detected through the minima event search with a window of size 100.

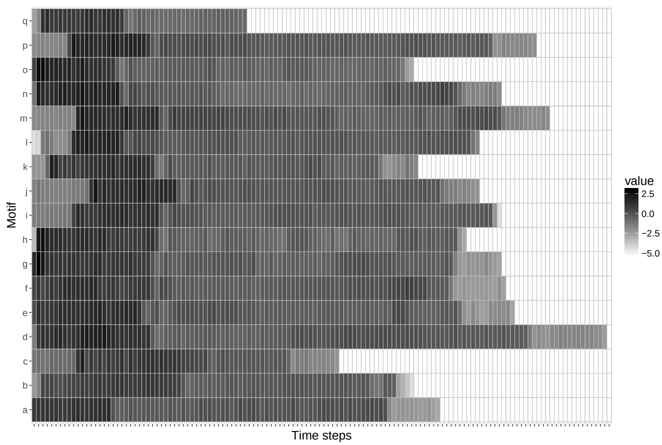


Figure 5: Heatmap of the individual subsequences found through the minima search. The subsequences start at the minimum and end at the next minimum. For a better comparison of their structure they are normalized ($\mu = 0, \sigma = 1$).

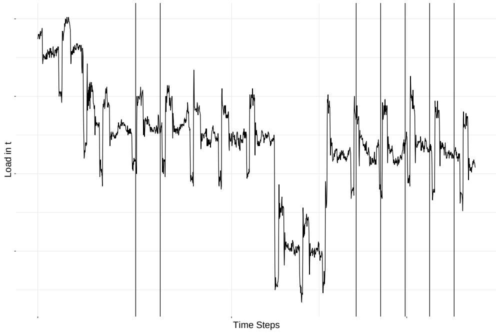


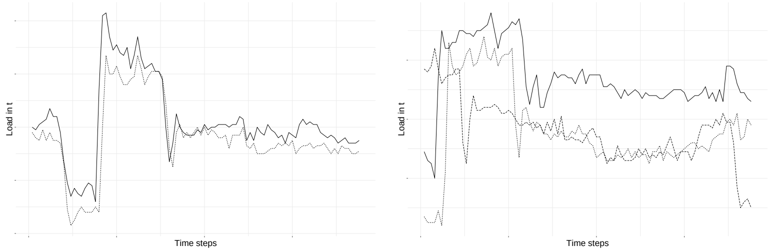
Figure 6: The steam consumption time series with the starting and ending point of the discovered motifs marked by the vertical lines.

process running would be the obvious choice for those parameters, which we do not have in our case. However, given the subsequences found in the event search, we can now use the mean time distance between the local minima found (w = 94). For the moment we do not allow for overlaps between the motifs.

The algorithm finds two different motifs, whose starting points in the time series can be seen in Figure 6. The individual instances of the motifs are displayed in Figure 7. Motif 1 occurs two times in the time series, while Motif 2 occurs three times. There are variations in the motifs, which is helpful for our flexibility examination later. However, they also do not look very different, which lets us assume they might be similar enough to be the same motif, if their comparison would start at the minimum as with the event search sequences.

To further investigate this idea, we compare the dtw distances of the motifs within themselves and in between each other. To be comparable, the motifs have all been normalized with mean $\mu = 0$ and standard deviation $\sigma = 1$. The result can be found in Table 1. The distances vary between 20.96 and 76.97, with the first one being the distance between the two instances in the first motif and the second one being the distance between the third instance of the second motif and the first instance of the first motif. The third instance of the second motif seems to be the odd one out.

This might be to the rather large drop at the beginning of the instance. The heatmap (Figure 8 confirms this idea, with the motif in the top row



(a) The first motif with its two instances (1a – 1b).

(b) The second motif and its three instances (2a – 2c).

Figure 7: The two motifs found by the motif discovery algorithm.

Table 1: The dynamic time warping distances for the motifs within themselves.

Motif	1a	1b	2a	2b	2c
1a	0	20.96	52.51	47.13	76.97
1b	20.96	0	49.04	42.25	75.07
2a	52.50	49.04	0	27.83	59.43
2b	47.13	42.26	27.83	0	59.63
2c	76.97	75.07	59.43	59.63	0
avg	39.51	37.46	37.76	35.37	54.22

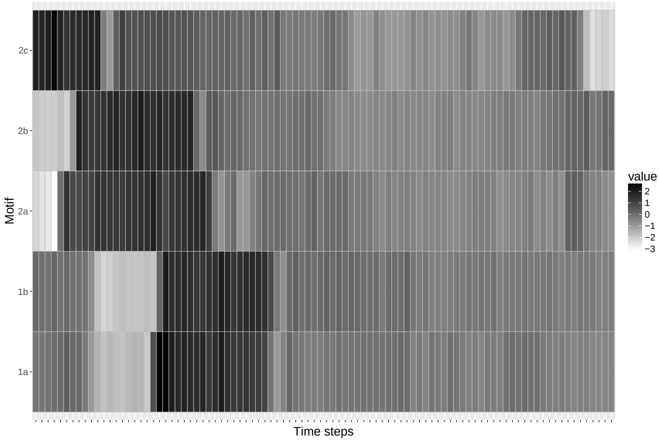


Figure 8: Heatmap of the two motifs found with their respective instances. The motifs are normalized with $\mu = 0$ and $\sigma = 1$.

(2c), having a shorter period at a high level energy intake before dropping. However, we will assume that the distance also for the third instance is small enough for our purposes and so assume that all five instances are triggered by the same mechanism.

As a next step we want to have a look at the subsequences from the event search. We have seen from the heatmap before, that they all exhibit a similar pattern but differ mainly in length (see Figure 5). We now compare the motifs found through the motif discovery algorithm with the subsequences determined through the event search. It is our aim to find how similar those motifs are to each other. Table 2 shows the difference calculated between each motif found with the motif discovery algorithm (1a – 2c) and the event search subsequences (a – q). We exclude subsequence g from further analysis as it is only five time steps long in comparison to the above 70 steps all other sequences are long.

As we can see in the table, the distances are for many sequences greater than those for the motifs within themselves. further describe dtw distances

Table 2: The dynamic time warping distances for each event search subsequence (a – q) with the motifs (1a – 2c).

Motif	a	b	c	d	e	f	g	h
1a	89.43	57.26	57.52	62.80	64.21	73.88	84.65	76.25
1b	93.45	62.16	61.59	65.89	63.81	78.00	94.78	78.00
2a	63.71	46.32	49.84	48.99	44.44	63.37	58.19	37.16
2b	70.06	60.01	58.86	62.49	62.87	74.96	74.75	45.24
2c	44.90	53.19	61.79	50.12	40.87	41.94	37.07	39.55
avg	72.31	55.79	57.92	58.06	55.24	66.43	69.89	55.24

Motif	i	j	k	l	m	n	o	p	q
1a	50.15	48.50	52.50	46.95	57.39	56.30	44.55	49.01	42.53
1b	53.47	52.31	53.78	46.72	58.73	59.77	57.35	50.18	49.43
2a	50.00	52.37	33.36	46.43	49.22	39.94	42.81	45.48	23.90
2b	42.31	44.16	37.70	32.58	42.19	33.34	52.59	44.18	32.43
2c	71.95	79.56	53.62	70.60	62.55	28.45	38.24	62.91	54.49
avg	53.57	55.38	46.19	48.65	54.02	43.56	47.11	50.35	40.56

3.2 Flexibility Measures

Given the subsequences we have found as which are similar enough to stem from the same underlying process, we now want to determine their flexibility. However, before going further into the analysis of the flexibility we can find by comparing the subsequences according to our flexibility measures, we first match the subsequences with the five motif instances. Thus, we will not take the motifs into our analysis if one of the subsequences already describe them.

Which subsequences are described already by the motifs?

Having established that the motifs are included in the subsequences ??, we now analyse the subsequences with the help of the above established flexibility measures. We first examine the energy flexibility of the motifs, before delving into the time flexibility of them.

3.2.1 Energy Flexibility

According to our three measures described before, we inspect the differences in energy usage from our subsequences. The results are displayed

Table 4: Characteristics of the found patterns in the steam demand time series.

Characteristic	Min	Median	Max	StdDev
Length	71.00	103.00	132.00	15.02
Intensity	293.94	294.37	390.97	10.31
Ramping Time	2.00			

in Table 4. The length of the processes varies between 71 and 132 time steps. However, the intensity for some motifs is rather similar although the lengths differ. This might indicate that no matter how long or short the process is, we need to use roughly the same amount of energy. The ramping time changes considerably between the instances. While the shortest ramping takes only two time steps, the longest takes more than update the time steps (exclude the negative ones).

Correlation between the length of the process and the intensity as measured by the area under the curve is $\rho = 0.8446$. Indicating that a longer process also uses more energy. However, ramping steps and the length of the process are only slightly positively correlated with $\rho = 0.2832$, as are the ramping steps and the intensity 0.3649.

Explain further

3.2.2 Time Flexibility

Finally, we consider the time flexibility of our sequences. Table 5 shows the hour and the day of the week where each local minimum is. As we can see, the weekdays and times of these possible starting points are well distributed throughout the week and day. This indicates that there is no dependency on the weekday or hour of the day to start the process. This might thus be a highly automated process or one with several shifts throughout the day. The fact that there is no clear pattern in the time could mean we are free to choose the starting point time-wise and thus have a great flexibility there. It could however also imply that the process highly depends on some other variables which we do not consider here.

3.3 Categorization of Flexibility

Best: Time Flexibility

Table 5: Starting date and time for each event.

Motif	Hour		Day of Week
a	9	am	Tuesday
b	0	am	Saturday
c	10	pm	Monday
d	9	am	Sunday
e	11	pm	Thursday
f	11	am	Tuesday
h	2	pm	Saturday
i	10	pm	Saturday
j	1	am	Thursday
k	1	pm	Monday
l	7	pm	Friday
m	11	am	Tuesday
n	5	pm	Saturday
o	3	pm	Thursday

Second Best: Ramping, Length

ToDo: Use correlations

4 Discussion

The flexibility we find in this paper is only *potential flexibility*. As we do not know anything about the process in terms of its technical properties or dependencies of other processes before or after it is running, we only indicate that there might be a potential as there has been some variety in the past. If this variety is only driven by technical features it would not be considered flexibility. Thus, in a next step to properly quantify the flexibility one has to talk to the process manager and let him rate the potentials found according to their usability.

The length of the motif varies thus all fixed window clustering approaches fail. One could run the algorithm several times with different window sizes but as we have no information about the processes possible lengths this approach seems tedious.

discuss results properly

5 Conclusion and Outlook

This paper proposed a new method to find motifs in chemical bath processes from steam consumption data.

In a next step we use sequence alignment to match the SAX representations of the subsequences and find how the motifs differ and are similar to each other instead of measuring the dynamic time warping distance. This would allow for a more detailed investigation into what can be used for flexibility purposes in the future. Furthermore, given the potential flexibility found we want to find market mechanisms to encourage the use of this flexibility by the process manager whenever this would be more efficient or cheaper and so forth

References

[1] E. Berlin and K. van Laerhoven. Detecting leisure activities with dense motif discovery. In Anind K. Dey, editor, *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 250–259, New York, NY, 2012. ACM.

[2] C. Cassisi, M. Aliotta, A. Cannata, P. Montalto, D. Patanè, A. Pulvirenti, and L. Spampinato. Motif discovery on seismic amplitude time series: The case study of mt etna 2011 eruptive activity. *Pure and Applied Geophysics*, 170(4):529–545, 2013.

[3] B. Chiu, E. Keogh, and S. Lonardi. Probabilistic discovery of time series motifs. In Ted Senator, editor, *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 493–498, New York, NY, 2003. ACM.

[4] L. Farinaccio and R. Zmeureanu. Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. *Energy and Buildings*, 30(3):245–259, 1999.

[5] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel. Disaggregated end-use energy sensing for the smart grid. *IEEE Pervasive Computing*, 10(1):28–39, 2011.

[6] E. F. Gomes, A. M. Jorge, and P. J. Azevedo. Classifying heart sounds using multiresolution time series motifs. In Bipin C. Desai, editor, *Proceedings of the International C* Conference on Computer Science and Software Engineering*, pages 23–30, New York, NY, 2013. ACM.

[7] M. Gupta, J. Gao, C. Aggarwal, and J. Han. Outlier detection for temporal data: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 26(9):2250–2267, 2014.

[8] C. Lin and C. L. Moodie. Hierarchical production planning for a modern steel manufacturing system. *International Journal of Production Research*, 27(4):613–628, 2007.

[9] A. Mueen, E. Keogh, Q. Zhu, S. Cash, and B. Westover. Exact discovery of time series motifs. In Haesun Park, editor, *Proceedings of the Ninth SIAM International Conference on Data Mining*, pages 473–484. SIAM, Philadelphia, 2009.

[10] M. Müller. *Information Retrieval for Music and Motion*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[11] B. Neupane, T. B. Pedersen, and B. Thiesson. Towards flexibility detection in device-level energy consumption. In Wei Lee Woon, editor, *Data analytics for renewable energy integration*, volume 8817 of *Lecture Notes in Computer Science*, pages 1–16. Springer, Cham, 2014.

[12] J. Serra, M. Müller, P. Grosche, and J. L. Arcos. Unsupervised detection of music boundaries by time series structure features. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[13] Y. Simmhan and M. U. Noor. Scalable prediction of energy consumption using incremental time series clustering. In Xiaohua Hu, editor, *IEEE International Conference on Big Data, 2013*, pages 29–36, Piscataway, NJ, 2013. IEEE.