High Frequency Non-intrusive Electric Device Detection and Diagnosis

Roman Jonetzko, Matthias Dettzler, Klaus-Uwe Gollmer, Achim Guldner, Marcel Huber, Rainer Michels and Stefan Naumann
Institute for Software Systems, Trier University of Applied Sciences, Environmental Campus Birkenfeld, Birkenfeld, Germany

Keywords: Pattern Recognition, Non-intrusive Load Monitoring, Classification, Fourier Descriptors.

Abstract: The number of electronic devices in households as well as in industrial workplaces is continuously growing because of progress in automation. Identifying unusual operating behavior, detecting device failures in advance, and recognizing energy saving potentials are key features to improve the reliability, safety, and profitability of these systems. Facing these tasks, today’s research is focused inter alia on a non-intrusive load monitoring approach, where the electrical signal is measured at a central point with modern hardware and processed by pattern recognition algorithms. Thus, we developed a smart meter prototype with a high sampling frequency, which allows for continuous measurement of the current and voltage from three-phase power lines. Besides this, in this paper we describe the usage of current-only measurement data (simple and safe installation using current transformers) with which we were able to classify state changes of a mobile air-conditioner with the help of Fourier descriptors as well as with additional voltage measurement.

1 INTRODUCTION

So far, currently available smart meter technology is used to measure inter alia electrical current, voltage, and power for determining the electrical energy consumption (or in-feed e.g. of photovoltaics) of a household or industrial building. For this purpose, it suffices to transfer the acquired data in intervals of seconds to several minutes (calculated from data measured at sampling rates in the kHz range). This usage is primarily focused on gaining knowledge about the immediate energy demand and to visualize the power curve. However, the analysis of signal data measured at high sampling rates implies another possible application of smart meters: the detection of switching events and the specific states of electric devices in a circuit by examining the signal for device specific patterns to perform load monitoring e.g. in technical work places.

In comparison to pure resistive devices such as light bulbs, many modern devices consist of electronic components which produce harmonics (integer multiples of the fundamental frequency) in the current signal, because of a nonlinear current-voltage behavior. With the calculation of these harmonics, which is possible by applying Fast Fourier Transformation to the measured signal, the frequency spectrum can be used as a “fingerprint” for different devices, as shown in figure 1. For our application we need to transfer measured data at appropriate speeds so that no events in the signal are missed. Because we did not find such a smart meter available at a reasonable price for consumers and small and medium enterprises on the market, we had to develop a smart meter for our application case.

![Figure 1: Current signal and amplitude spectrum of an LED-lamp (a), compared to a compact fluorescent bulb (b).](image)

A cost-efficient method for load monitoring is to install a smart meter at a central point (e.g. distribution...
box). This removes the need for measuring every device individually but also makes disaggregation algorithms necessary to detect individual devices from the measured sum signal. Because no intrusion into machines or devices is needed, this method is called non-intrusive appliance load monitoring (NIALM). The aim of load monitoring is to get knowledge about specific devices and their operation states, improve fault detection and identify energy efficiency potentials. In residential applications, this allows for identifying devices, in industrial applications with complex machines, it allows for identifying machine components.

In our paper we first consider the state-of-the-art in NIALM research, focus in the following chapter on our smart meter prototype. Afterwards, we describe appropriate features for classification, show inter alia how to use Fourier descriptors for classification with current measurement data only, and how we realized the classification of the active devices.

2 RELATED WORK

Using high-frequency sampled data for NIALM algorithms was already mentioned by (Leeb et al., 1995), where they describe transient event detection via spectral envelopes. More recent research used up to date high frequency measurement hardware available at the market (high-investment), which directly outputs calculated complex Fourier coefficients (FC) and fed this data into neural networks (Srinivasan and Liew, 2006). They indeed reached good detection-rates, however they had no variable loads present in their measurements. This presence of variable loads is a widely stated problem for the algorithms, which is described in various publications, e.g. (Lee et al., 2005). With the finding of strong correlations between higher harmonics and the active power Lee et al. subtracted the estimated power of the variable load from the sum signal, showing that this applies if a single variable load is present. (Zoha et al., 2012) gives an overall overview of steady-state (using features like power change) and transient-state (using features like start-up current transients) disaggregation methods and points out that it is still a challenge for researchers to develop a solution which is able to detect all kinds of device types. A promising approach could be an implementation of a multi-feature/multi algorithm solution, which (Liang et al., 2010) propose, with the extraction of additional features, such as the complete waveform of a cycle period and instantaneous admittance.

So far, little research has been conducted on low-investment non-intrusive measuring devices. In (Guldner et al., 2013) we presented a centralized, low investment data acquisition prototype based upon a modified consumer energy meter by Reichelt Electronics Co. In this paper we further developed the hardware (cf. section 3) and software (cf. section 5).

3 SMART METER HARDWARE AND DATA PROCESSING

The purpose of the smart meter in our case is not to use it for utility billing but rather for performing load monitoring. Therefore no exact active power calculation but transferring measured data in real-time is important. As mentioned in the introduction, we developed a smart meter to match these requirements, which includes the following parts:

- Low-cost Measurement Instruments. Current transformers and voltage transformers, which make direct access to the power lines unnecessary (direct connection to the power lines can be used for voltage measurement, too)
- Poly Phase Measurement Chip. An ATME 90E36A chip reads current and voltage signals of up to three phases at sampling rate of 7.324 kHz
- Real-time Processing of Measured Data. To match the requirement of real-time processing, we use a Teensy 3.1 microcontroller development system that features a MK20DX256 ARM Cortex-M4 with 72 MHz, 64 kB RAM memory, a SPI and an USB Port
- Wireless Data Transfer to a Server. A Raspberry Pi transfers the data via wi-fi, which provides galvanic isolation between potentially expensive server hardware and the measured power lines of up to 400 volts (in case of direct connection to the power lines).

Figure 2 depicts the system structure. The data generated by the measurement chip is gathered by the Teensy. Via the SPI bus, the Teensy sends the chip (in slave-mode) into direct memory access mode (DMA). Then - in master-mode - the chip sends the read current and voltage measurements directly over the SPI bus, where the Teensy collects them into two buffers of 2,048 values. One buffer is filled with the alternating current and voltage readings of each phase, while the second buffer is propagated to data preprocessing.

For synchronizing the data, the ATMEL chip provides a zero-crossing signal for each phase via a zero-crossing detection pin, which propagates the start and finish of a period in the current or voltage signals.
This pin is also read by the Teensy in an interrupt, which switches the data collection to the other buffer after the maximum number of full signal periods that fit within the 2,048 value limit have been read (in German power lines with 50 Hz, this results in six periods being read). This approach fixes the measuring window to a full multiple of a current- and voltage period, allowing us to simply apply the Fast Fourier Transform (FFT) to the signal as-is, without the need for a window function and without distorting the spectral estimate.

For processing the data, at first, a complex, in-place, fixed-point Radix-4 algorithm is used to calculate the FFT. Afterwards, the complex values of the FC are extracted and the fundamental and first 40 harmonics selected. Thus, the data can be read continuously (no fluctuation in the current and voltage signals is missed) and is reduced from 8,192 bytes to 656 bytes (real and imaginary part of the 40 harmonics plus the fundamental).

Optionally - if only the current signal is measured - Fourier descriptors (description see chapter 4.2) are applied. Finally, the data is sent to a Raspberry Pi (or any other computer) over the USB-Port of the Teensy. There, the data can be transferred to a database server or directly be visualized, disaggregated or clustered.

Figure 2: Data acquisition and processing in the devised smart meter board.

Figure 3: Amplitude curve of the FC envelope corresponding to 50 Hz signal component.

4 CLASSIFICATION FEATURES

Previous research showed that so far, it is implausible to detect all kinds of electric devices through their electrical signal using only a single feature (Zeifman and Roth, 2011). The reason is a different load behavior of different kinds of devices that cause the extraction of specific features needed to detect devices of specific load behaviors. Therefore, the calculation of FC of the measured raw signal data (at high sampling rate), provides information content about signals shape and phasing. In the following, a short overview is given of the appliance classifications stated by (Hart, 1992) which is kept up by subsequent NIALM research. Afterwards, the calculation of FC and their subsequent processing is detailed.

Permanent Devices. Devices which are always active and having a constant load (e.g. hard-wired alarms, phones, routers).

On-off Devices. Devices with only one active state, not operated constantly, and having constant load (e.g. light bulbs, electric motors without speed control).

Finite State Machines. Devices with more than one active state in between which can be switched and having a constant load in each state (e.g. dryer, multi state kitchen devices)

Continuously Variable Devices. Devices which can have a continuously variable load behavior, having no stepped variation of load (e.g. speed controlled electric motors like in power drills, modern energy efficient pumps and fans).

To depict the differences, the amplitude curve of the FC envelope of the different device types is shown in figure 3.
4.1 Choice of Fourier Coefficients and Influence of Noise

It is widely known that, under certain assumptions, the Fourier series of a signal \( s = s(t) \) with period \( T \) and fundamental frequency \( f_0 = 1/T \) can be rewritten as its Fourier series. When a signal is in a steady state, in our case, we can assume that the fundamental frequency is (nearly) 50 Hz, so that \( T = 0.02s \) (since \( s = s(t) \) is given by the power supply). Because the observation time is given by a rectangular window \( w \), the result of the fast Fourier transform is an approximation of the Fourier Transform of \( s \cdot w \) (we denote the Fourier Transform of \( s \) as \( \mathcal{F}(s) \)). If the length/time of \( w \) is chosen in a way that \( T \mid w \), meaning \( w = k \cdot T \) with \( k \in \mathbb{Z} \), one can see, with the choice of \( \omega = \omega_m = m \cdot 2\pi f_0 \), where \( m \in \mathbb{N}_0 \) that \( \mathcal{F}(s \cdot w)(\omega_m) \) gives the calculation formula for the (theoretical) \( m \)-th complex FC (noted as \( c_m = c_m(s) \)) of a \( T \)-periodic signal \( s = s(t) \). Since the signal \( s \) is real (current, voltage, or power), it is also known that \( c_{-m} = c_m \). In steady state, by using the FFT, we compute \( \mathcal{F}(s \cdot w)(\omega) \), and choosing \( \omega = m \cdot 2\pi f_0 \) \( (m \in \mathbb{N}_0) \), we receive approximately the (complex) FC of the Fourier series of \( s \). Since there are fluctuations in \( f_0 \) (meaning that the fundamental frequency is not stable at \( f_0 = 50 \text{ Hz} \) permanently), in practice, we can find a good approximation to these complex FC. In experiments we observed a variation of fundamental frequency from the power supply in a range between 49.8 Hz and 50.2 Hz.

If the measured signal (the observation time is given by window \( w \)) is nearly in steady state, these complex FC can be used to characterize this signal. Using these FC, we can calculate a finite Fourier series which gives a good approximation to the observed signal. These considerations motivate our choice of feature selection: In the DFT-Calculation (DFT: Discrete Fourier transform), we use these DFT-values with \( \omega \), so that it is approximately \( \omega = m \cdot 2\pi f_0 \), where \( m \in \mathbb{N}_0 \). Using this approach, we also have two restrictions to observe: First, the number of interesting FC is bounded by the Nyquist-Shannon sampling theorem. In the discrete form, we observe this constraint also by the conjugate symmetry of a DFT calculated vector. The second constraint is a limitation given by the magnitude of the noise.

4.2 Fourier Descriptors

As mentioned in chapter 3, our system allows us to classify the devices, measuring only the current signal (cf. chapter 5.1). In this case a calculation of Fourier descriptors is required, which the following passage describes.

No Capturing of the Phase Shift Between Current and Voltage. When the measurement of the current signal \( s = s(t) \) is started at a random point without measuring the voltage at the same time, we have no knowledge about the phasing of the signal. Having a continuously changing window position of the Fourier Transform, this causes a continuous shift in the real and imaginary parts of FC \( c_k = c_k(s) \), which is not feasible for device detection. Calculating the amplitude spectrum, we can get a constant set of frequency components of the signal, usable as a “fingerprint” of a device. But it lacks information about any-angle relation of the frequency components. Although we have no knowledge about the phasing of the fundamental frequency related to the voltage signal, we can relate the phasing of high-order-harmonics of the current signal to the fundamental frequency of the current signal to get a description about the signal shape, which represents additional information for device detection. Calculation of Fourier descriptors \( c_k = c_k(s) \) (cf. formula (1)) provides this and makes evaluation of periodicity of the signal possible:

If \( c_1 \neq 0 \):
\[
\begin{align*}
c_1 &= |c_1| \cdot e^{i\varphi}; \quad \varphi \in [0, 2\pi) \\
c_2 &= \frac{c_2}{(e^{i\varphi})^2}; \quad c_3 = \frac{c_3}{(e^{i\varphi})^3}; \ldots; \quad c_k = \frac{c_k}{(e^{i\varphi})^k}
\end{align*}
\]

\( k \) : (positive) harmonic index

Figure 4: Illustration of the first three FC before (l.) and after applying Fourier descriptors (r.).

For two signals \( s_1, s_2 \) we have:
\[
\tilde{c}_k(s_1 + s_2) \neq \tilde{c}_k(s_1) + \tilde{c}_k(s_2)
\]

Since Fourier descriptors are not additive (cf. formula (2)) all possible states/combinations of measured devices have to exist as reference patterns. To show the additional information acquired by using Fourier descriptors in comparison to the amplitude spectrum, we generated two signals of different
4.3 Ratios Between nth and 1st Harmonic

The problem of detecting devices when variable loads are present caused us to research features which remain constant for a variable load during their operation. As a result, the ratio between high-order harmonics and the fundamental frequency seems to be relatively constant in initial tests (cf. figure 6). However, the ratio is not an additive feature, so sum signals with more than one active device cannot be used. For generating a signal where only one device is active, the delta-form has to be calculated. The delta-form is the difference between the measured instantaneous data and the data at the previous measurement point. In this period of time between two measured data, we assume that only one device is switched.

Figure 6: The load variation of the hotplate (a); the variation in amplitude spectrum (b); ratios diagram (c), where a good stability of ratio between 3rd and 1st FC is recognizable (150 Hz to 50 Hz component ratio of around 20%).

(Höck, 2009) states that devices of different electric function have specific values of those ratios. In Table 2 the classification of devices used in our experiments is listed.

5 CLASSIFICATION

To perform state and device detection from measured signals, we used three different pattern recognition approaches. We performed real time recognition of device states of a mobile air-conditioner with a clustering algorithm. Here we especially tested classifying with current only measurement data, and also tested artificial neural networks. The used setup is described in section 5.1. With the setup described in section 5.2 we measured combinations of different devices and applied a nearest-neighbor classifier to that historic dataset.

5.1 Benchmark Setup 1

In this setup we analyzed a mobile air-conditioner individually considering its different operation states, which are specified in Table 1. This means the focus is not on detecting the single components but detecting the different states of the device. Therefore, the new Teensy smart meter was used to benefit from its continuous measurement method.

5.1.1 Applying Cluster Analysis

For creating clusters we collected the first 17 FC in a row vector. We measured 2,000 datasets for each possible state (corresponds to approx. 4 minutes per state). These historic datasets of all possible states
Table 2: Overview of Ratios for Measured Devices and Physical according to (Höck, 2009).

<table>
<thead>
<tr>
<th>Device</th>
<th>Ratio $n^{th}$ to 1$^{st}$ harmonic [%]</th>
<th>Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 2</td>
<td>n = 3</td>
</tr>
<tr>
<td>Air-conditioner fan</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Speakers</td>
<td>35</td>
<td>18</td>
</tr>
<tr>
<td>Pedestal fan</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Compact fluorescent lamp (CFL) without ECG</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Halogen lamp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixer</td>
<td>62</td>
<td>25</td>
</tr>
<tr>
<td>Hotplate</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Food processor</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>Electric knife</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1: Overview of the mobile air-conditioner states.

<table>
<thead>
<tr>
<th>State</th>
<th>Active device components</th>
</tr>
</thead>
<tbody>
<tr>
<td>FanLow</td>
<td>Evaporator fan performs air-circulation in three different power states (42 W, 44 W, 51 W)</td>
</tr>
<tr>
<td>FanMed</td>
<td>Condenser fan (74 W), condenser water pump (9 W), evaporator fan (42 W, 44 W, 51 W), compressor (500...550 W)</td>
</tr>
<tr>
<td>CoolHigh</td>
<td>Condenser fan (74 W), condenser water pump (9 W), evaporator fan (42 W, 44 W, 51 W), compressor (500...550 W)</td>
</tr>
</tbody>
</table>

are averaged (by components) and the calculated $C^{17}$ average is used as center point for the respective cluster. For calculating cluster radii, as a first approach we used the following formula:

$$r = \frac{Q_{0.25} + Q_{0.75}}{2} + 2 \cdot IQR$$  \hspace{1cm} (3)

$$Q :$$ quartile  
$$IQR :$$ interquartile range

Figure 7 shows the data flow of applying the cluster algorithm to measured input data. If a new measured vector is located in more than one cluster, it is assigned to the cluster with the minimum distance to its cluster center. Is the vector not located in any cluster, it is defined as an outlier, if it falls below the 25% quartile (low-outlier) or exceeds the 75% quartile (high outlier), respectively by more than 2 times the IQR.

The detection of the three fan only switching states (low/med/high) with the clustering algorithm using only current measurement data (no capture of phase shift of the signal, thus applying Fourier descriptors) works nearly as well as with phase capture (cp. table 3). In comparison to that, using only amplitudes results in a much lower detection rate. During compressor and evaporator fan states are not recognizable by the clustering algorithm, even when phase shift of the signal is considered through zero-crossing detection of voltage signal. The high difference in load scale between compressor and evaporator fan makes detecting small changes of the fan’s power consumption difficult, a problem we also observed in setup 2.

Table 3: Detection rates applying cluster algorithm for three fan states and different preprocessed measurement data.

<table>
<thead>
<tr>
<th>Measurement data</th>
<th>Detection rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current and voltage (phase capture)</td>
<td>93%</td>
</tr>
<tr>
<td>Current only (Fourier descriptors applied to)</td>
<td>82%</td>
</tr>
<tr>
<td>Current only (amplitudes)</td>
<td>39%</td>
</tr>
</tbody>
</table>

5.1.2 Applying Artificial Neural Networks

Besides the clustering algorithm, we tested if an artificial neural network (ANN), which it is another appropriate method to solve pattern recognition problems, is able to detect all states of the mobile air-conditioner. Using the MathWorks Matlab software, we chose a two-layer feed-forward network with sigmoid transfer function and back propagation training and used as input vector the result of a principle component analysis (PCA) of the FC, which represents a lower number of weighted and sorted features. The target-vector contains whether or not the particular device is active (element contains 1) or inactive (element contains 0).

The trained ANN yields the same result as the clustering algorithm: it was also able to correctly classify the fan states, but not able to distinguish between cooling states due to the high differences in power consumption.

5.2 Benchmark Setup 2

The second setup consists of nine (cf. table 4) devices which were measured individually for creating reference patterns and combined with an automated...
device switching system composed of power distribution units (Energenie EG-PMS2, multi sockets which are switchable by PC through a USB connection). The goal of this setup was to disaggregate the single devices from a measured sum signal. For this measurement, in an early project phase, we used the ZigBee smart meter prototype which measures at 4 kHz sampling rate current and voltage signal and does not provide continuous measurement.

Table 4: Classification of the devices via their load classes.

<table>
<thead>
<tr>
<th>Device</th>
<th>Nominal Power [W]</th>
<th>Load Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air conditioner fan</td>
<td>75</td>
<td>on-off</td>
</tr>
<tr>
<td>Speakers</td>
<td>3</td>
<td>on-off</td>
</tr>
<tr>
<td>Pedestal fan</td>
<td>30</td>
<td>on-off</td>
</tr>
<tr>
<td>CFL</td>
<td>25</td>
<td>on-off</td>
</tr>
<tr>
<td>Halogen lamp</td>
<td>30</td>
<td>on-off</td>
</tr>
<tr>
<td>Mixer</td>
<td>40</td>
<td>on-off</td>
</tr>
<tr>
<td>Hotplate</td>
<td>15</td>
<td>cont. variable</td>
</tr>
<tr>
<td>Food processor</td>
<td>130</td>
<td>cont. variable</td>
</tr>
<tr>
<td>Electric knife</td>
<td>65</td>
<td>on-off</td>
</tr>
</tbody>
</table>

5.2.1 Applying Nearest-neighbor Classifier

To detect individual loads from the aggregated signals, a nearest-neighbor classifier has been implemented. At first, the current signals of the switchable distribution unit’s microcontrollers were subtracted from the measured reference dataset (individual devices) and combination dataset. Thereby, a disturbance in the calculation of combinations from the individual signals could be avoided. Using FC, calculated through discrete Fourier transformation of the measured raw signals, were classified with a brute-force algorithm, which means all possible combinations were calculated from the reference pattern (individual measurements of all existing devices) and the sum of the features was compared to the feature of the measured dataset signal (real measurement of combination). The calculated combination with the smallest Euclidean distance to the real measured combination was classified as the current combination.

Applying the brute force algorithm using nearest neighbor classifier to the complete dataset of 9 devices (containing devices from different load behavior classes) using the first 39 FC as features a relatively low recognition rate of 27% for this dataset was computed (cf. table 5). Separating out combinations of devices which include devices with load behavior different than on-off-devices significantly increased the detection rate.

Table 5: Detection rates of complete and reduced dataset.

<table>
<thead>
<tr>
<th>Left out devices</th>
<th>None</th>
<th>Speaker, hotplate, mixing machine</th>
</tr>
</thead>
<tbody>
<tr>
<td>correctly classified combinations</td>
<td>139</td>
<td>38</td>
</tr>
<tr>
<td>Total no. of combinations</td>
<td>511</td>
<td>64</td>
</tr>
<tr>
<td>Percentage</td>
<td>27%</td>
<td>91%</td>
</tr>
</tbody>
</table>

We identified the appearance of variable loads as primary cause for confounding the NIALM algorithm. The devices with variable loads in this dataset are the hotplate where a temperature controller adjusts the current demand and the food processor which
runs up to higher rotational speed meaning that current demand decreases (whereby the efficiency factor increases). A continuously variable load coincides with continuous changes of the frequency spectrum. This implies that the averaged values of FC cause a high similarity of different compared combinations. In section 4.3, we presented an appropriate feature (ratios between FC) that is likely to solve this problem. Because we often switched more than one device at the same time during the measurement of the dataset, we are not able to use this feature yet, because it needs the delta-form (only one device may switch at the same time).

In comparison with setup 1, we could observe the same problem in device detection, when there is a large difference between the load scale of present devices. Variations of the power consumptions of high load devices can exceed smaller devices power consumption level. Here, the speakers load is as high as the variation range of power of the mixing machine. So far, we did not find any feature which could solve this.

6 SUMMARY AND OUTLOOK

Our developed smart meter prototype, which is currently in testing phase, is able to provide information about the electric signal without missing any event occurring during time of measurement. We pointed out that with Fourier descriptors, classification without consideration of the voltage signal is possible for observing state changes of finite state machines. For specific machines which have a finite state load behavior and whose states are of a similar load scale, this could be a low cost and safe method of supervising. Like previous research, we observed the problem of variable loads for the disaggregation algorithms at setup 2. Therefore, additional extracted features like ratios between FC could help to identify the existence of variable loads and will be tested in further setups. Another approach is to improve transient signal detection, where envelopes of FC during device start-up and shut-down could be allocated with specific devices. The goal is to make all information contained in the FC available and combining features extracted from short time windows and long time windows in detection algorithms.

ACKNOWLEDGEMENTS

This paper evolved from the research and development project “Process Monitoring and Improved Energy Efficiency of Technical Workplaces via Smart Meters” (PEBiS), which is sponsored by the foundation “Stiftung Rheinland-Pfalz für Innovation” grant number 961-386261/1048. The contents of this document are the sole responsibility of the authors and can under no circumstances be regarded as reflecting the position of the foundation “Stiftung Rheinland-Pfalz für Innovation”.

REFERENCES


