# Adaptive Load Signature Coding for Electrical Appliance Monitoring over Low-Bandwidth Communication Channels

Andreas Reinhardt
Technische Universität Clausthal
Clausthal-Zellerfeld, Germany
reinhardt@ieee.org

***Abstract*—Collecting and analyzing power consumption data from electrical appliances is a key enabling element for grid-related services, e.g., load forecasting or anomaly detection. Device-level sensors (*smart plugs*) have found widespread use to collect such data. However, they prevalently report an electrical appliance's power consumption at a rate of one reading per second in order to limit the resultant communication traffic. With mains voltage frequencies of 50/60 Hz, undersampling and the consequent loss of spectral information result from the use of such reporting rates. Moreover, as most smart plugs only report real power consumption values, important supplementary features (e.g., the phase shift between voltage and current or the magnitude of reactive power) are not available when using such devices. In this work we present a data processing system design that exploits the recurring nature of electrical current waveforms in order to facilitate the provision of data at a high resolution whilst keeping the corresponding data rate requirements low. Our design, called ALSCEAM, is applicable to voltage and current waveforms collected at high sampling rates, thus spectral components are implicitly included in collected traces. Instead of transferring raw readings to external processing services, however, local data processing routines are being employed to detect and eliminate redundancies. Thus, a high data fidelity is maintained while network traffic is reduced by more than 95% in many cases. All functionalities are implemented in a proof-of-concept system design and evaluated in practice.**

***Index Terms*—Load Signature Collection; High-Frequency Sampling; Adaptive Load Signature Coding**

## I. Introduction

Since the publication of George Hart's seminal work on non-intrusive load monitoring (NILM) in 1985 [1], research activities to extract and utilize the information content in electricity consumption data have seen an almost exponential increase. A key insight of *energy analytics* is that power consumption patterns are suitable indicators to infer information on both appliance and user activities in residential, commercial, and industrial settings. Many data processing algorithms have consequently been presented to analyze energy consumption data, e.g., to characterize household types [2, 3], infer building occupancy [4], or predict the future use of electricity [5, 6]. Their designs often also highlight one of the principal limitations of NILM [7, 8]: It relies on the usage of a single measurement point (usually a smart electricity meter). Thus, differentiating between appliances of the same type operated in within the same dwelling is complicated.

In order to overcome this limitation, deployment strategies for sensors to capture electricity consumption in a more fine-grained fashion have emerged. For example, the use of higher sampling rates has shown to improve the potential of electricity consumption data processing [7, 9]. Supplementally, more sensing points can be deployed in buildings, either by installing separate meters into each electrical circuit [10, 11], or by inserting them into the power cords of all relevant devices [12]. Unfortunately, such device-level sensors (*smart plugs*) often report values only once per second, thus transients and spikes of shorter duration commonly remain undetected. Furthermore, appliances with inductive or capacitive components incur a reactive power consumption, whose detection requires the synchronous sampling of voltage and current. As most smart plugs are not fitted out with voltage sensors, they are unable to differentiate between real and reactive power consumption. At last, mains voltage and current consumption waveforms rarely resemble perfect sinusoids. Instead, slight distortions are ubiquitous due to the wide presence of non-linear loads, e.g., switch-mode power supplies. Quantifying such deviations from the ideal sinusoidal shape is, however, rarely supported by smart plugs.

Aforementioned shortcomings of commercially available devices have motivated researchers to design embedded sensing systems to capture and transmit electricity consumption data. The key distinctive property of such systems is the rate and resolution at which they provide data. Some solutions offer raw samples without any prior processing, e.g., *ACme* [13], and consequently generate significant network traffic. Others (e.g., *WCSN* [14]) apply lossy data processing algorithms and only report characteristic values, e.g., RMS current and crest factor, at time intervals on the order of seconds or even minutes. Less traffic is being generated, as voltage and current waveforms are omitted from transmission, yet the potential of in-depth analyses is narrowed at the same time. We overcome these restrictions of existing solutions by presenting an adaptive data coding scheme named ALSCEAM. It enables accurate energy analytics by providing consumption data at high temporal resolutions while minimizing bandwidth requirements.

This paper is structured as follows: We motivate the need for the adaptive encoding of energy consumption data by means of a practical example in Sec. II. In Sec. III, we

978-3-901882-99-9 

present ALSCEAM and elaborate on the design decisions taken. We demonstrate ALSCEAM's efficacy based on traces from an existing data corpus as well as measurements collected ourselves in Sec. IV. Existing works related to our contribution are discussed in Sec. V before we draw conclusions in Sec. VI.

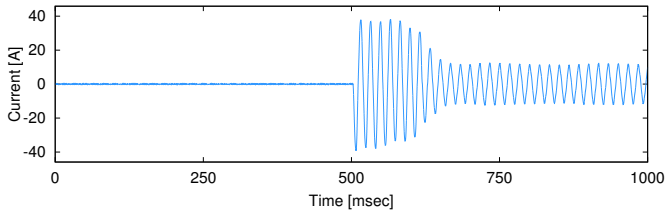
## II. Load Signatures

The *load signature* models an electrical appliance's characteristic power consumption during the course of its operation. It is defined by the current flowing through the device as well as the voltage across its terminals. The analysis of load signatures has been extensively studied in energy analytics research, e.g., in [17–19], with the objective of extracting high-level information from collected data. One key insight gained in the studies was the correlation between the sampling frequency at which load signatures are collected and their information content [9, 20–22]. Two principal types of load signatures have thus been defined by Zeifman and Roth in [23], which we briefly revisit as follows to cater for a common understanding.

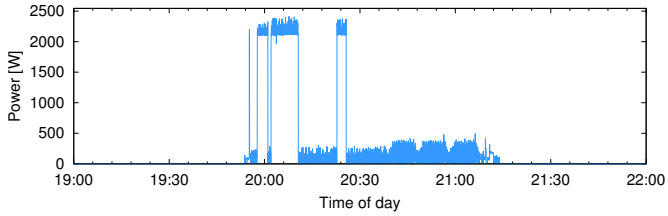
### A. Load Signature Types

Many of today's electrical appliances are non-linear devices. That is, their power consumptions contain frequency components that exceed the mains frequency by far. Their load signatures must consequently be captured using sampling frequencies of several kilohertz [13, 24] to retain such characteristics, which in turn enable services like the detection of an appliance's mode of operation or its mechanical wear [25]. Load signatures that rely on voltage and current signals captured at frequencies much higher than the frequency of the mains voltage are commonly referred to as *microscopic* load signatures.

PLAID [15] is a data set that contains microscopic appliance inrush signatures collected at 30 kHz sampling rate, i.e., 500 times the mains frequency. A sample load signature from the PLAID data set is shown in Fig. 1a. It shows a washing machine's current consumption in the moment the appliance is switched on. The sinusoidal nature of the current consumption waveform is clearly visible, as well as the fact that the power consumption in the initial mains periods after activation is largely different than the comparably steady waveform afterwards. While this level of detail appears promising for data analytics, a key downside of microscopic load signatures is their size. Even when applying lossless compression, data collected at 24 bits resolution and 16 kHz sampling rate sum up to almost 5 gigabytes per appliance per day [26].




(a) Microscopic load signature of a washing machine's activation current.




(b) Macroscopic load signature of a washing machine's power consumption.

Fig. 1. Sample traces of macroscopic and microscopic load signatures of washing machine appliances (traces from PLAID [15] and Tracebase [16]).

Alternatively, load signatures can be composed of averaged readings of a load's power consumption, reported at low frequencies, such as once per second. This lowered resolution makes it impossible to reconstruct the actual shape of the appliance's current intake. However, such *macroscopic* load signatures cater to the limitations imposed by wireless communications and storage systems. Even though the level of detail achievable from microscopic load signatures is no longer available, macroscopic load signatures still allow for the analysis of long-term trends and patterns on time scales of minutes, hours, or even days. An example for this type of load signature is shown in Fig. 1b, which shows the power consumption of a washing machine from the Tracebase data set [16], in which data has been collected at 1 Hz rate. The washing machine's heating cycles with power consumptions in excess of 2,000 W can be easily distinguished from rinsing phases with considerably lower power demand. Even lower sampling rates, e.g., once per minute or once in every 15-minute interval, are used in practice.

The conversion of a microscopic load signature into its macroscopic representation is easily possible, yet a lossy process. For example, the multiplication of the current waveform shown in Fig. 1a with the corresponding voltage samples (also part of the data set) yields the washing machine's power consumption which can then be integrated over the course of one second. The three largely different phases visible in Fig. 1a (off, initial inrush, steady state), however, all take place within one second. Consequently, they can no longer be discerned in the macroscopic load signature once the data have been converted, despite their potential relevance for load signature analysis algorithms.

### B. Load Signature Compression

While there is unequivocal agreement among researchers on the enormous information content of microscopic load signatures, their significant storage requirements are equally widely acknowledged. Several gigabytes of data are generated per day when collecting raw (i.e., unprocessed) microscopic load signature data [26–28]. Lossless data compression algorithms can mitigate this transmission and storage requirement by eliminating redundancies in the data. The efficacy when applying data compression to macroscopic load signatures has been studied in detail, e.g., in [29, 30]. However, their applicability to microscopic signatures has not yet been assessed, despite the anticipated high compressibility of microscopic load signatures because of the recurrent nature of current and voltage waveforms.

TABLE I
ACHIEVABLE COMPRESSION GAINS WHEN ENCODING THE PLAID TRACE OF A LAPTOP COMPUTER.

| Data format | Algorithm | Input | Size (bytes) | Savings | Input | Size (bytes) | Savings | Input | Size (bytes) | Savings |
|---|---|---|---|---|---|---|---|---|---|---|
| CSV (ASCII) | *none* | 1/60 second = 1 period (500 samples) | 6,439 | — | 2 seconds = 120 periods (60,000 samples) | 769,327 | — | 1 minute (1,800,000 samples) | 22,268,520 | — |
| | ZIP | | 2,351 | 63.5 % | | 233,286 | 69.7 % | | 152,968 | 99.3 % |
| | BZIP2 | | 1,935 | 69.9 % | | 152,057 | 80.2 % | | 166,827 | 99.2 % |
| IEEE 754 (32-bit) | *none* | | 4,000 | — | | 480,000 | — | | 13,920,000 | — |
| | ZIP | | 2,523 | 36.9 % | | 236,737 | 50.7 % | | 83,737 | 99.4 % |
| | BZIP2 | | 2,719 | 32.0 % | | 152,057 | 68.3 % | | 98,299 | 99.3 % |
| WAVE (16-bit) | *none* | | 2,044 | — | | 240,044 | — | | 6,960,044 | — |
| | ZIP | | 2,085 | -2.0 % | | 180,021 | 25.0 % | | 39,342 | 99.4 % |
| | BZIP2 | | 2,097 | -2.6 % | | 153,961 | 35.9 % | | 37,585 | 99.5 % |
| | FLAC | | 1,205 | 41.0 % | | 123,250 | 48.6 % | | 3,487,478 | 49.9 % |
| WAVE (32-bit) | *none* | | 4,044 | — | | 480,044 | — | | 13,920,044 | — |
| | ZIP | | 2,794 | 30.9 % | | 251,656 | 47.6 % | | 84,048 | 99.4 % |
| | BZIP2 | | 3,429 | 15.2 % | | 184,086 | 61.7 % | | 118,236 | 99.1 % |
| | FLAC | | 1,811 | 55.2 % | | 193,978 | 59.6 % | | 5,740,332 | 58.8 % |

By way of example, we have analyzed the compression gains when a microscopic load signature taken from the PLAID data set [15] (file `29.csv`; laptop computer) is supplied to different compression methods. The following four data encodings have been used to prepare the input data before applying the compression algorithms:

- ASCII (numerical values for current and voltage, comma-separated), for example: `1.38,159.93`
- IEEE 754 floating-point numbers at single precision (i.e., 32 bits each per voltage and current sample)
- RIFF WAVE (PCM data) at 32 bits per sample
- RIFF WAVE (PCM data) at 16 bits per sample

After transcoding the input file into each of these formats, we have used the ZIP[1] and bzip2[2] tools to compress the resulting output files. Moreover, we have encoded the two latter file formats into the Free Lossless Audio Codec (FLAC) format using SoX[3]. Table I shows the resulting output sizes for three input traces of different lengths ($1/60$ s, 2 s, and 60 s) as well as the achievable savings with regard to the data's corresponding uncompressed representations. It is notable that across all configurations, measurable savings are already achievable by transcoding the ASCII data into a format with fixed-size entries. However, compression gains vastly differ depending on the size of the input data. When only a single mains period is being considered and encoded as 16-bit WAVE format, its compressed representation even grows larger than the uncompressed file. On the contrary, input waveforms with longer durations lead to highly compacted data compression results, as visible in the rightmost column of Table I.

While these results may appear promising for the compression of microscopic load signatures, their practical use suffers from three limitations. Firstly, our analysis has shown the highest compression gains when large volumes of input data have been compressed. Hardware restrictions of embedded measurement devices, however, often disallow for the buffering of correspondingly large data fragments before applying compression. As soon as shorter fragments (e.g., only one period of the mains voltage) are compressed, however, less savings are observed and the introduced overhead for header information may even increase the output size. Secondly, embedded systems to capture load signatures do not necessarily feature the computational capabilities to execute the listed compression algorithms due to their demand for memory and computational power. For example, running the bzip2 algorithm requires at least 400 kB of RAM according to its documentation. Lastly and most importantly, data compression may mitigate the requirements to the bandwidth of the communication channel across which the data are transferred. However, their actual processing, i.e., the analysis of the waveforms for patterns of interest and the removal of irrelevant data, still remains to be performed on external systems with sufficient computational capabilities.

[1]Zip 3.0 by Info-ZIP, Release: 5 July 2008
[2]Version 1.0.6, Release: 6 September 2010
[3]SoX v14.4.2, Release: 22 February 2015

## III. ADAPTIVE LOAD SIGNATURE CODING

Finding a sensible trade-off between data collection and data reporting is of critical importance and our primary motivation for the work presented in this paper. The objective is to provide data at high resolution (i.e., in the fashion of microscopic load signatures) when notable changes occur, yet alleviate the resultant requirement for bandwidth and/or storage by omitting redundant data. To accomplish this goal, we investigate the potential of using load signatures with adaptive sampling rates. Whenever significant changes occur on microscopic level, fine-grained detail about the appliance's current consumption shall be reported. In turn, during an appliance's steady state of operation, it suffices to only collect macroscopic features in order to reduce transmission and storage requirements.

### A. Exploiting Load Signature Periodicities

Virtually all power grids worldwide run on alternating current (AC) at mains frequencies of 60 Hz or 50 Hz. This gives microscopic load signatures a distinctive property: Due to the

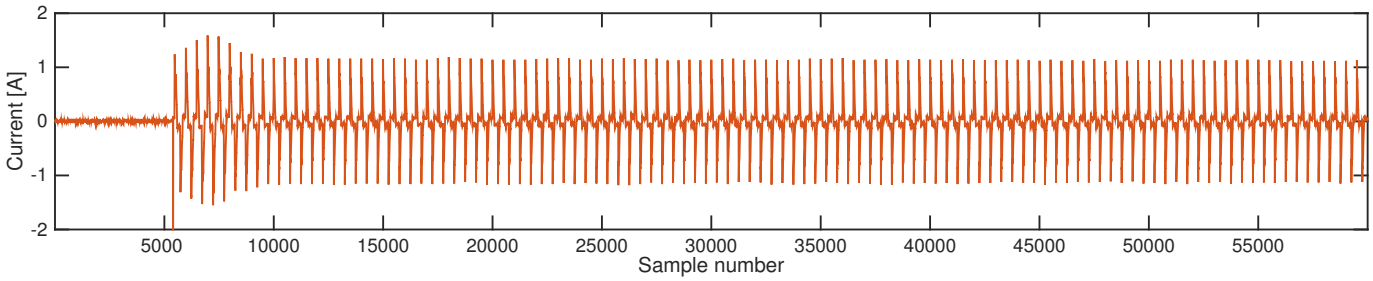



Fig. 2. Compact fluorescent lamp trace from the PLAID data set [15] (`3.csv`). The negative peak overshoot at sampling point 5,399 has been omitted.

sinusoidal waveform of the voltage, recurring patterns can be found in appliances' current consumptions as well. This is visible in Fig. 2, which shows a two-second long excerpt of the current consumption waveform of a compact fluorescent lamp. After an initialization phase (from samples 5,400 to 9,400), a perceptibly periodic continuation of the waveform can be observed. Considering the short duration of the plotted trace, we can assume the conditions of the appliance's environment as well as its internal state to be mostly static during the data collection. A recurrent current consumption behavior is consequently expected. However, corresponding redundancies are rarely omitted by devices that collect microscopic load signatures (cf. Sec. V). As a result, bandwidth requirements can arise that exceed the capabilities of low-power wireless communication links.

We tackle this need for bandwidth by proposing our adaptive load signature coding scheme called ALSCEAM. In contrast to general-purpose data compression algorithms (such as the ones discussed in Sec. II-B), ALSCEAM takes the semantics of load signatures into consideration. That is, it builds on the observation that voltage and current waveforms generally exhibit recurring patterns. Deviations from previously observed current consumption can often be attributed to changes in the underlying appliance's mode of operation. Thus, corresponding microscopic load signatures must be transmitted in a lossless manner, given their importance for successful load signature analyses. In contrast, repeating current consumption patterns with identical or highly similar waveforms often carry little information content and can be omitted to reduce the data volume.

In order to accomplish this adaptive data transmission mechanism, ALSCEAM performs a local valuation of current consumption waveforms. It relies on a configurable similarity threshold parameter to this end, which allows it to determine whether two current consumption waveforms can be considered to be similar. In essence, ALSCEAM reduces data traffic by returning *hybrid* load signatures: It reports microscopic load signatures in a lossless fashion when changes have been observed, while resorting to macroscopic load signature reporting during phases of steady current flows.

### B. ALSCEAM's Operation

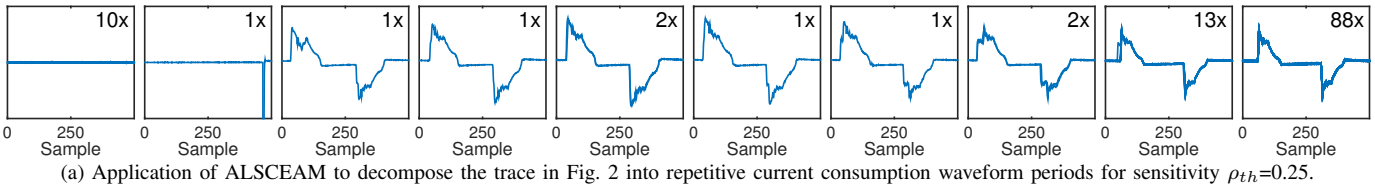
To optimally exploit the periodicity of microscopic load signatures, ALSCEAM needs to process data fragments of exactly one period length (20 ms in power grids operated at 50 Hz and $16.\overline{6}$ ms at 60 Hz). ALSCEAM relies on a windowing algorithm to this end, which segments the continuous waveforms into data periods based on the detection of the voltage signal's zero-crossings. In case of the PLAID data set [15], a sampling rate of 30 kHz has been used for appliances on a 60 Hz power grid, thus 500 samples are present in each window. Once a full window a samples has been collected, the data processing routines are invoked.

Algorithm 1 describes ALSCEAM's data processing operation. It relies on two buffer arrays $x$ and $\bar{x}$ of identical sizes, each of which contains one full period of current consumption data. Buffer $\bar{x}$ is used to store the last waveform that has been reported in full, i.e., as a microscopic load signature. In turn, $x$ is updated whenever a new current consumption waveform has become available, e.g., from an attached current sensor (line 2). As soon as new data has arrived, the dissimilarity between $x$ and $\bar{x}$ is computed (line 3) and compared to the user-definable sensitivity threshold $\rho_{th}$ (line 4). If the deviation between $x$ and $\bar{x}$ is too large, the microscopic waveform is transmitted without further processing (line 5), and the contents of $\bar{x}$ are updated accordingly (line 6). Moreover, the scaling factor $\varphi$ is re-computed (line 7), as will be discussed below. In case the deviation between $x$ and $\bar{x}$ is below the sensitivity threshold $\rho_{th}$, only features relevant to capture a macroscopic load signature (currently only the root mean square of the current) are being reported.

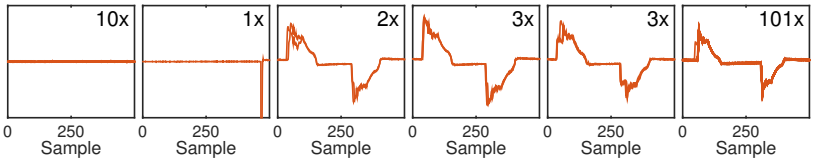
**Algorithm 1** Operation of ALSCEAM

**Variables**: $x$, $\bar{x}$ (arrays with capacity to store one period of the appliance's current waveform); $\varphi$ (scaling factor)
**Parameter**: $\rho_{th}$ (sensitivity threshold)
**Initialization**: $x \leftarrow \varnothing$; $\bar{x} \leftarrow \varnothing$; $\varphi \leftarrow 0$

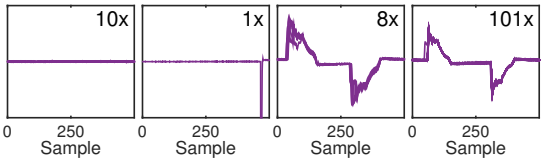
1: **loop**
2: x ← next current waveform period
3: compute $\rho(x, \bar{x})$ according to Eq. (1)
4: **if** $\rho(x, \bar{x}) \geq \rho_{th}$ **then**
5: report microscopic load signature
6: $\bar{x} \leftarrow x$
7: update $\varphi$ according to Eq. (2)
8: **else**
9: compute and report macroscopic load signature
10: **end if**
11: **end loop**




(a) Application of ALSCEAM to decompose the trace in Fig. 2 into repetitive current consumption waveform periods for sensitivity $\rho_{th}$=0.25.




(b) Decomposition of the trace with a less strict similarity requirement ($\rho_{th}$=0.45).




(c) Decomposition of the trace with $\rho_{th}$=0.6.

Fig. 3. Application of ALSCEAM to separate the trace shown in Fig. 2 into microscopic load signatures. The number of occurrences of the same microscopic signature is annotated in the top-right corner of each diagram.

We compute the deviation $\rho$ between $x$ and $\bar{x}$ by means of the root mean square error (RMSE), as given in Eq. (1), where $N$ is the length of the sample. Other metrics to determine the similarity of waveform periods can be integrated into ALSCEAM easily.

$$\rho(x,\bar{x}) = \varphi \cdot \sqrt{\frac{\sum_{t=0}^{N-1}\left(x(t)-\bar{x}(t)\right)^2}{N}} \tag{1}$$

Subsequently, $\rho$ is weighted by a scaling factor $\varphi$. This factor is chosen proportionally to the maximum encountered amplitude range in $\bar{x}$, and computed according to Eq. (2). The scaling factor is required to ensure that $\rho_{th}$ can be specified relative to the amplitude swing of $\bar{x}$.

$$\varphi = \begin{cases} \max\bar{x} - \min\bar{x}, & \text{if } (\max\bar{x} - \min\bar{x}) > 1 \\ 1, & \text{otherwise} \end{cases} \tag{2}$$

### C. Study of a Sample Trace

To demonstrate how the choice of parameter $\rho_{th}$ affects the number of required microscopic data transmissions, we have conducted a preliminary sensitivity analysis for $\rho_{th} = [0.25, 0.45, 0.6]$. We have used the 120 periods of the compact fluorescent lamp's current consumption, as shown in Fig. 2, as the input data to ALSCEAM. The resulting microscopic load signature transmissions, as well as their total number of occurrences (denoted in the top-right corner) are given in Figs. 3a to 3c. Note that all signatures that meet the similarity requirement are superimposed in the diagrams to show their deviations more closely.

Across all considered traces, ten periods of idleness are observed initially. Only the first occurrence is transmitted as a microscopic load signature, whereas the remaining nine repetitions are reported in macroscopic form. The initial inrush spike visible in the second subfigure strongly deviates from its preceding periods of inactivity. It is thus also transmitted as a microscopic load signature for all analyzed values of $\rho_{th}$. During the further course of the waveform, differences are observed depending on the setting of the sensitivity value. When being more tolerant to deviations (e.g., for $\rho_{th}$=0.6, visible in Fig. 3c), only two more microscopic load signatures are transmitted for the remaining 109 mains periods. In contrast, when ALSCEAM is set up to be less tolerant to deviations, six (for $\rho_{th}$=0.45) or even ten (for $\rho_{th}$=0.25) microscopic load signatures result. In summary, the number of microscopic load signatures transferred in full can be successfully reduced from 120 to 10, 6, or 4. This is equal to savings between 91.7% and 96.7% in terms of microscopic load signatures. In practice, the savings are slightly lower because ALSCEAM substitutes redundant microscopic load signatures by a macroscopic value.

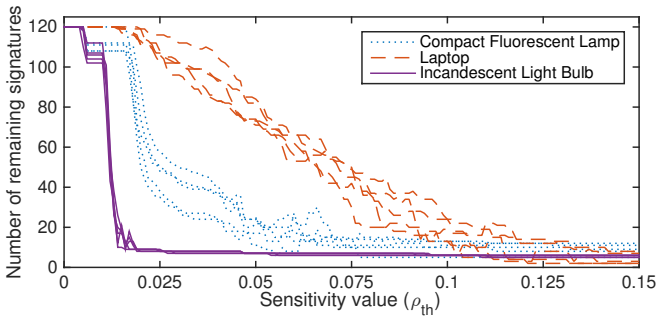
## IV. Evaluation

The transmission of microscopic load signatures incurs a significant communication overhead. By reducing the number of microscopic signatures that require transmission, this bandwidth requirement can be effectively lowered. However, the resultant reduced data resolution may hamper the operation of load signature analysis algorithms at the same time. It is thus important to understand the relation between achievable savings and retained accuracy. The trade-off between these two parameters is governed by the choice of $\rho_{th}$. To guide users of ALSCEAM in its selection process, we present insights from two analyses in this section.

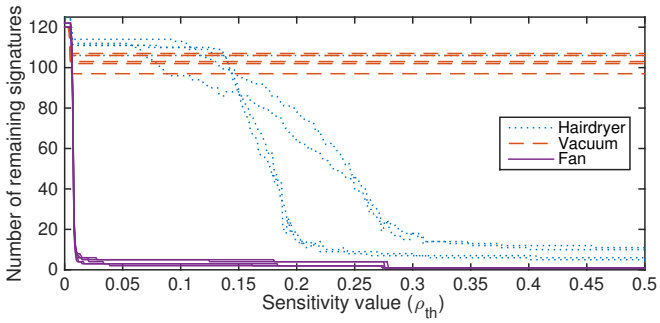
### A. Size Reductions for "Off-On" Transitions

In a first experiment, we analyze how $\rho_{th}$ is related to the type of appliance the load signature monitor is attached to. To this end, we rely on device activation traces from the PLAID data set [15]. All selected traces were of 2 seconds duration, and of type "off-on," i.e., traces during which the given appliance has been switched on. We have selected nine appliance types from the data set, considered five "off-on" traces for each of them, and evaluated the impact of $\rho_{th}$ on the number of microscopic signatures to be transmitted. Results for our simulations are provided in Fig. 4, in which we have grouped appliances based on the scale of their sensitivity values: 0–0.15 in Fig. 4a, 0–0.5 in Fig. 4b, and a range from 0–1 in Fig. 4c.

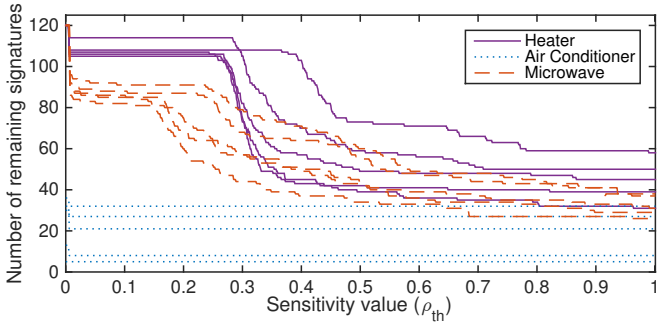
The diagrams show strong variations in the efficacy of ALSCEAM when applied to different appliances. The subset of appliances compared in Fig. 4a show highly similar recurring current waveforms. Thus, even when choosing $\rho_{th}$=0.1,




(a) Sensitivity value vs. appliance type (first batch of appliances).




(b) Sensitivity value vs. appliance type (second batch of appliances).




(c) Sensitivity value vs. appliance type (third batch of appliances).

Fig. 4. Relation between the sensitivity value $\rho_{th}$ and the resulting number of microscopic load signatures for different appliance types.

large reductions are achievable. The same is true for the fan in Fig. 4b and the air conditioner in Fig. 4c. However, other appliances impose limitations on the applicability of ALSCEAM. As visible in Fig. 4c, the numbers of microscopic load signatures for heater and microwave oven only experience reductions by approximately 50%. After a slight initial drop, the number of microscopic signatures transferred for the vacuum cleaner only drops slowly and almost linearly (cf. Fig. 4b). In fact, even for $\rho_{th}$=10, ALSCEAM requires 28 microscopic load signature transmissions to fully capture its current consumption.

It needs to be noted at this stage, however, that we have deliberately considered "off-on" traces in this evaluation. Thus, it can be expected that the number of signatures transmitted is higher than during regular operation. Moreover, the choice of a different similarity metric (cf. Sec. III-B) in conjunction with corresponding sensitivity parameter values, can also be expected to directly influence the number of microscopic signatures that need to be transmitted in full.

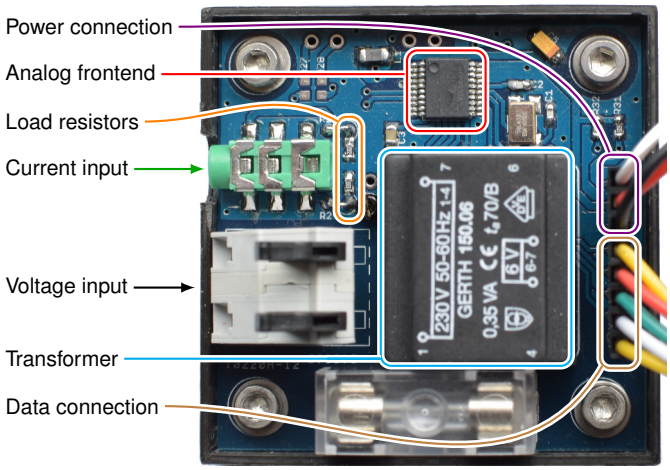



Fig. 5. Photo of the TUCap board used to collect practical measurements.

### B. Size Reductions during Steady-state Operation

Supplementary to the analysis of device activation traces taken from PLAID, we have run tests on ALSCEAM's efficacy during steady-state operation. To this end, we use the TUCap measurement board [31], which uses a Microchip MCP3910 analog frontend to synchronously sample voltage and current flows. Readings are collected at 36 kHz sampling frequency and a resolution of 16 bits per sample. It is burdened with a series of load resistors configured to fully utilize the frontend's input range for a $\pm$433 V voltage range and $\pm$16 A primary current. A photography of the board is shown in Fig. 5. In order to allow for the collection and transmission of load signatures, we connect the analog frontend to a Teensy 3.0 system [32], which features sufficient computational power to execute ALSCEAM. A voltage zero-crossing detection routine triggers its execution whenever 720 samples, i.e., one mains period at 50 Hz, were collected.

We have sequentially attached four appliances to the TUCap board and evaluated ALSCEAM's compression gains for different values for $\rho_{th}$ when capturing 5,000 mains periods (i.e., 100 seconds) each. The resulting number of microscopic load signature transmissions are tabulated in Table II. Reductions of the number of microscopic load signature transmissions can be attained in all cases, even for very small sensitivity values. In fact, for the monitor and the printer, less than 2% of the collected microscopic load signatures needed to be transmitted in full when $\rho_{th}$=0.01.

In a final experiment, we have connected the data collection system running ALSCEAM to a laser printer during the course

TABLE II
NUMBER OF MICROSCOPIC LOAD SIGNATURE TRANSMISSIONS DURING STEADY-STATE APPLIANCE OPERATION.

| | Selected sensitivity value ($\rho_{th}$) | | | | | |
|---|---|---|---|---|---|---|
| Appliance | 0.0 | 0.005 | 0.01 | 0.015 | 0.02 | 0.025 |
| Battery charger | 5,000 | 502 | 336 | 303 | 288 | 200 |
| Printer (standby) | 5,000 | 271 | 4 | 2 | 2 | 2 |
| Laptop computer | 5,000 | 4,109 | 2,029 | 1,452 | 722 | 456 |
| LCD Monitor | 5,000 | 940 | 82 | 4 | 2 | 2 |

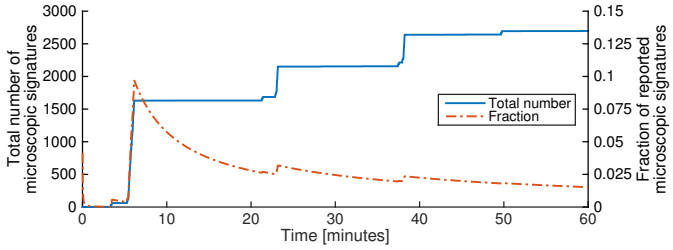



Fig. 6. Microscopic signatures reported throughout the long-term study.

of 60 minutes. The parameter $\rho_{th}$ has been configured to a value of 0.15, as per the insights shown in Table II where this value led to measurable reductions, yet not the omission of all data. The objective of this experiment was to get a long-term impression of the number of microscopic load signatures reported by ALSCEAM. The results are visualized in Fig. 6; the total number of transmitted are signatures shown as a continuous line therein. During the printer's activity (it was printing five pages each at minutes 5, 23, 38), an increase in the number of transmitted signatures can be observed. However, only very few updated signatures were generated during phases of inactivity. The right-hand side y-axis visualizes the fraction of signature with regards to the total number of mains voltage cycles. In total, after one hour of its operation, only 2697 out of the 180,000 hourly mains periods were transmitted; i.e., savings of 98.5% were achieved in practice.

## V. Related Work

The limited usability of commercially available smart plugs have motivated many researchers to design platforms to measure electrical current flows. Resulting devices include Plug [24], ACme [13], SmartMeter.KOM [33], WCSN [14], and YoMo [34]. While these devices are capable of sampling data at higher temporal resolutions than their commercially available counterparts, their data processing functionalities are often limited. In fact, research platforms either forward data to a collection device without any prior processing (i.e., they report data at the native sampling rate), or apply lossy data processing algorithms to return characteristic values (e.g., RMS current, crest factor, etc) at the 1 Hz interval prevailing among commercial platforms. The former approach, however, results in an enormous bandwidth requirement, whereas the latter solution disallows for the detection of short-term fluctuations. Given the small resource footprint of ALSCEAM, retrofitting existing high-resolution sensing devices should easily be possible. It would enable them to output hybrid load signatures, i.e., to provide microscopic detail of collected waveforms where meaningful, and resort to macroscopic data reporting for the remaining time. Apart from TUCap [31] (cf. Sec. IV-B), we are currently not aware of any other embedded sensing device that outputs hybrid load signatures.

Alternative means to reduce the traffic volume on embedded sensing systems have been presented in the area of data compression. A range of lossless data compression algorithms were proposed with specific adaptations to accommodate the lossy channel characteristics prevalent in networks of embedded systems. Algorithms in this category mostly include fault-tolerant variants of established data compression algorithms, e.g., RT-LZW [35] and FT-AHC [36], which cater to lossy channels by relying on retransmissions where necessary. A commonality of these algorithms is their independence of the type of input data. However, their generic definition does not allow them to take characteristics and/or semantics of the input data into consideration. A variant of BZIP2 for embedded systems, named SBZIP, is presented in [37]. While this implementation would in theory allow the system to achieve compression gains similar to those listed in Table I, its application domain is the efficient transfer of firmware updates. As firmware images only need decompression on a sensing system, SBZIP is confined to provide decompression routines with no option to compress data as well.

At last, works on the combination of both domains have emerged which specifically considered to compression of electricity consumption data [29, 38]. These approaches specifically take the characteristics and semantics of consumption data into account, thus higher compression gains are possible than when using generic data compression algorithms. However, as they have only been used to compress macroscopic load signature data so far, no temporal bounds were imposed on their operation. Observed data processing times have been shown to range between 50–350 milliseconds [29]. Microscopic load signature processing, however, necessitates reaction times on the order of $16.\overline{6}$–20 ms (cf. Sec. III-B), i.e., whenever the waveform of a full mains period has been captured. ALSCEAM has specifically been designed to operate on data sampled at rates of several kilohertz and fulfills this criterion. It can thus be easily accommodated on embedded data collecting systems.

## VI. Conclusions

Load signature analysis research has received significant attention in the past decade. A key contributing factor for the wide interest in this area is the increasing availability of energy and power consumption data collected by embedded sensing devices. A strong limitation of most commercially available platforms, however, is their limited data reporting rate. When data is being transmitted once per second or even less often, only macroscopic load signatures can be created, i.e., traces that omit specific features of an appliance's current consumption waveform. To overcome this limitation, we have presented ALSCEAM, an algorithm that adaptively encodes microscopic load signatures to retain high information content while reducing the traffic volume. It is configurable to application needs: By choosing small values for $\rho_{th}$, highly accurate consumption traces can be recorded. In turn, selecting a larger value for the parameter reduces generated traffic and may thus enable ALSCEAM's use in scenarios where bandwidth is scarce. In both simulation studies and practical experiments, coding gains in excess of 90% could be observed while maintaining the fidelity of microscopic load signatures where necessary.

## References


[1] G. W. Hart, "Prototype Nonintrusive Appliance Load Monitor," MIT Energy Laboratory and Electric Power Research Institute, Tech. Rep., 1985.

[2] M. Weiss, A. Helfenstein, F. Mattern, and T. Staake, "Leveraging Smart Meter Data to Recognize Home Appliances," in *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2012.

[3] C. Beckel, L. Sadamori, T. Staake, and S. Santini, "Revealing Household Characteristics from Smart Meter Data," *Elsevier Energy*, vol. 78, 2014.

[4] W. Kleiminger, C. Beckel, and S. Santini, "Household Occupancy Monitoring Using Electricity Meters," in *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, 2015.

[5] V. Bianco, O. Manca, and S. Nardini, "Electricity Consumption Forecasting in Italy using Linear Regression Models," *Energy*, vol. 34, no. 9, 2009.

[6] A. Reinhardt, D. Christin, and S. S. Kanhere, "Can Smart Plugs Predict Electric Power Consumption? A Case Study," in *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous)*, 2014.

[7] M. Gulati, S. S. Ram, and A. Singh, "An in Depth Study into Using EMI Signatures for Appliance Identification," in *Proceedings of the ACM Conference on Embedded Systems for Energy-Efficient Buildings (BuildSys)*, 2014.

[8] M. Kahl, A. Ul Haq, T. Kriechbaumer, and H.-A. Jacobsen, "WHITED – A Worldwide Household and Industry Transient Energy Data Set," in *Proceedings of the 3rd International Workshop on Non-Intrusive Load Monitoring (NILM)*, 2016.

[9] S. Tomkins, J. Pujara, and L. Getoor, "Disambiguating Energy Disaggregation: A Collective Probabilistic Approach," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2017.

[10] A. Marchiori, D. Hakkarinen, Q. Han, and L. Earle, "Circuit-Level Load Monitoring for Household Energy Management," *IEEE Pervasive Computing*, vol. 10, no. 1, 2010.

[11] D. Jung and A. Savvides, "Estimating Building Consumption Breakdowns Using ON/OFF State Sensing and Incremental Sub-meter Deployment," in *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2010.

[12] T. W. Hnat, V. Srinivasan, J. Lu, T. I. Sookoor, R. Dawson, J. Stankovic, and K. Whitehouse, "The Hitchhiker's Guide to Successful Residential Sensing Deployments," in *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2011.

[13] X. Jiang, S. Dawson-Haggerty, P. Dutta, and D. E. Culler, "Design and Implementation of a High-Fidelity AC Metering Network," in *Proceedings of the 8th International Conference on Information Processing in Sensor Networks (IPSN)*, 2009.

[14] D. Porcarelli, D. Balsamo, D. Brunelli, and G. Paci, "Perpetual and Low-cost Power Meter for Monitoring Residential and Industrial Appliances," in *Proceedings of the Design, Automation Test in Europe Conference Exhibition (DATE)*, 2013.

[15] J. Gao, S. Giri, E. C. Kara, and M. Bergés, "PLAID: A Public Dataset of High-resolution Electrical Appliance Measurements for Load Identification Research: Demo Abstract," in *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings (BuildSys)*, 2014.

[16] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz, "On the Accuracy of Appliance Identification Based on Distributed Load Metering Data," in *Proceedings of the 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT)*, 2012.

[17] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, "Load Signature Study — Part I: Basic Concept, Structure, and Methodology," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, 2010.

[18] H. H. Chang, K. L. Chen, Y. P. Tsai, and W. J. Lee, "A New Measurement Method for Power Signatures of Nonintrusive Demand Monitoring and Load Identification," *IEEE Transactions on Industry Applications*, vol. 48, no. 2, 2012.

[19] L. Du, D. He, R. G. Harley, and T. G. Habetler, "Electric Load Classification by Binary Voltage-Current Trajectory Mapping," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, 2016.

[20] C. Efthymiou and G. Kalogridis, "Smart Grid Privacy via Anonymization of Smart Metering Data," in *Proceedings of the 1st IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2010.

[21] M. G. Tauber, F. Skopik, T. Bleier, and D. Hutchison, "A Self-organising Approach for Smart Meter Communication Systems," in *Proceedings of the 7th IFIP TC 6 International Workshop on Self-Organizing Systems (IWSOS)*, 2014.

[22] A. Reinhardt, F. Englert, and D. Christin, "Averting the Privacy Risks of Smart Metering by Local Data Preprocessing," *Pervasive and Mobile Computing*, vol. 16, 2015.

[23] M. Zeifman and K. Roth, "Nonintrusive Appliance Load Monitoring: Review and Outlook," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 1, 2011.

[24] J. Lifton, M. Feldmeier, Y. Ono, C. Lewis, and J. Paradiso, "A Platform for Ubiquitous Sensor Deployment in Occupational and Domestic Environments," in *Proceedings of the 6th International Symposium on Information Processing in Sensor Networks (IPSN)*, 2007.

[25] U. A. Orji, Z. Remscrim, C. Laughman, S. B. Leeb, W. Wichakool, C. Schantz, R. Cox, J. Paris, J. L. Kirtley Jr., and L. K. Norford, "Fault Detection and Diagnostics for Non-intrusive Monitoring Using Motor Harmonics," in *Proceedings of the 25th Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2010.

[26] J. Kelly and W. Knottenbelt, "The UK-DALE Dataset: Domestic Appliance-level Electricity Demand and Whole-house Demand from Five UK Homes," *Scientific Data*, vol. 2, no. 150007, 2015, available for download at http://jack-kelly.com/data/.

[27] H. Maaß, H. K. Cakmak, F. Bach, R. Mikut, A. Harrabi, W. Süß, W. Jakob, K.-U. Stucky, U. G. Kühnapfel, and V. Hagenmeyer, "Data Processing of High-Rate Low-Voltage Distribution Grid Recordings for Smart Grid Monitoring and Analysis," *EURASIP Journal on Advances in Signal Processing*, no. 1, 2015.

[28] A. Ul Haq, T. Kriechbaumer, M. Kahl, and H. A. Jacobsen, "CLEAR: A Circuit Level Electric Appliance Radar for the Electric Cabinet," in *Proceedings of the IEEE International Conference on Industrial Technology (ICIT)*, 2017.

[29] M. Ringwelski, C. Renner, A. Reinhardt, A. Weigel, and V. Turau, "The Hitchhiker's Guide to Choosing the Compression Algorithm for Your Smart Meter Data," in *Proceedings of the 2nd IEEE Conference and Exhibition / ICT for Energy Symposium (ENERGYCON)*, 2012.

[30] A. Unterweger, D. Engel, and M. Ringwelski, "The Effect of Data Granularity on Load Data Compression," in *Proceedings of the 4th D-A-CH Conference on Energy Informatics (EI)*, 2015.

[31] A. Reinhardt, "TUCap: A Sensing System to Capture and Process Appliance Power Consumption in Smart Spaces," in *Proceedings of the 16th GI/ITG KuVS Fachgespräch "Sensornetze" (FGSN)*, 2017.

[32] PJRC LLC., "Teensy USB Development Board," available online at https://www.pjrc.com/teensy/index.html, last access on 8 November 2017.

[33] A. Reinhardt, D. Burkhardt, P. S. Mogre, M. Zaheer, and R. Steinmetz, "SmartMeter.KOM: A Low-cost Wireless Sensor for Distributed Power Metering," in *Proceedings of the 6th IEEE Workshop on Practical Issues in Building Sensor Network Applications (SenseApp)*, 2011.

[34] C. Klemenjak, D. Egarter, and W. Elmenreich, "YoMo: The Arduino-based Smart Metering Board," *Computer Science – Research and Development*, vol. 31, no. 1, 2016.

[35] C. M. Sadler and M. Martonosi, "Data Compression Algorithms for Energy-Constrained Devices in Delay Tolerant Networks," in *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems (SenSys)*, 2006.

[36] A. Guitton, N. Trigoni, and S. Helmer, "Fault-Tolerant Compression Algorithms for Delay-Sensitive Sensor Networks with Unreliable Links," in *Proceedings of the 4th IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2008.

[37] N. Tsiftes, A. Dunkels, and T. Voigt, "Efficient Sensor Network Reprogramming through Compression of Executable Modules," in *Proceedings of the 5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, 2008.

[38] A. Unterweger and D. Engel, "Resumable Load Data Compression in Smart Grids," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, 2015.