Electric Appliance Classification Based on Distributed High Resolution Current Sensing

Andreas Reinhardt, Dominic Burkhardt, Manzil Zaheer, Ralf Steinmetz
Multimedia Communications Lab, Technische Universität Darmstadt, Germany
{andreas.reinhardt, dominic.burkhardt, manzil.zaheer, ralf.steinmetz}@kom.tu-darmstadt.de

***Abstract*—Today's solutions to inform residents about their electricity consumption are mostly confined to displaying aggregate readings collected at meter level. A reliable identification of appliances that require disproportionate amounts of energy for their operation is generally unsupported by these systems, or at least requires significant manual configuration efforts. We address this challenge by placing low-cost measurement and actuation units into the mains connection of appliances. The distributed sensors capture the current flow of individual appliances at a sampling rate of 1.6kHz and apply local signal processing to the readings in order to extract characteristic fingerprints. These fingerprints are communicated wirelessly to the evaluation server, thus keeping the required airtime and energy demand of the transmission low. The evaluation server employs machine learning techniques and caters for the actual classification of attached electric appliances based on their fingerprints, enabling the correlation of consumption data and the appliance identity. Our evaluation is based on more than 3,000 current consumption fingerprints, which we have captured for a range of household appliances. The results indicate that a high accuracy is achieved when locally extracted current consumption fingerprints are used to classify appliances.**

I. Introduction

Rising prices for electric energy on a global scale have lead to an increased awareness of how energy is spent in households and office spaces. The identification of unnecessary consumption is however complicated by the fact that electricity bills are distributed in intervals of months to years only. Being confined to a posteriori analyses, the end user can only detect that an excessive demand for energy has occurred during a billing period, but not determine its origin. This limitation of traditional metering is partially overcome by the introduction of smart meters [1], which cater to the availability of near real-time consumption data. As a result, users can identify periods of high electric power demand faster.

Determining the actual appliances that have contributed to the high power demand, however, lies beyond the capabilities of smart meters, because the metering is usually performed at the building level. The *disaggregation* of load curves collected by smart meters is thus an essential function, as it allows the end user to determine which appliance is contributing to the total power demand. Numerous load disaggregation approaches (e.g., [2], [3], [4]) exist, but most of them are confined to a subset of appliances that meet special criteria; the majority of approaches can only detect appliances correctly if their power draw exceeds 150 watts and they have defined states of operation, between which repeatable transitions occur.

In contrast to the aforementioned centralized load identification approaches, less constraints are imposed on the monitored appliance when distributed monitoring devices are used. The Plugwise system [5] is based on placing sensors into the mains connection of each appliance to be monitored. Each Plugwise device returns averaged power consumptions over periods of one hour each at a resolution of 0.2 watts. Therewith, the operation cycles of major loads can be extracted and their contribution to the total demand can be calculated on a hourly basis. Linking the deployed Plugwise devices to the identity of the metered appliance however remains a manual task, which can quickly become cumbersome when a large number of appliances is present.

In this paper, we advance this current state of the art by presenting a solution that infers the type of an electric appliance based on its power consumption data. The primary motivation for our work is to eliminate the configuration effort required to manually link the distributed power sensors to the type of the attached appliance. Instead, our architecture permits the classification of appliances in a plug-and-play fashion; it autonomously collects the required current measurements, extracts a characteristic fingerprint from the data collected, and forwards this identification feature to an evaluation engine, returning a classification result which can be used to annotate the data. The contributions of this paper are as follows:

- We introduce the architecture as well as implementation details of our data collection system, which is based on an extension of our SmartMeter.KOM platform [6], an embedded sensing system comprised of a current sensor, a microcontroller, and a wireless communication module.
- We discuss the data preprocessing taking place locally on the metering nodes, as well as details about the machine learning implementation used in our backend system.
- We present the evaluation results of three practical experiments, which are based on more than 3,000 collected current waveform traces.

First, we provide an overview of related work in Sec. II. An introduction of our appliance classification system architecture is given in Sec. III, in which we also present the local preprocessing mechanisms which extract characteristic fingerprints from device inrush and steady-state currents in a microscopic (i.e., in the time scale of milliseconds) manner in more detail. We evaluate the classification accuracy of our approach in Sec. IV, and conclude this paper in Sec. V.

978-1-4673-2129-7/12/$31.00 ©2012 IEEE

II. Related Work

First approaches to identifiy electric appliances based on their load signatures have emerged in the 1980s under the name of Non-Intrusive Appliance Load Monitoring (*NIALM*) [2]. The basic idea behind NIALM was to place a current sensor between the electricity meter and its socket in order to capture a household's total power draw at a high temporal resolution. The NIALM system extracts transients from the observed power traces and matches them against the characteristics of previously trained appliance types. While the approach has been shown to be capable of correctly identifying a small number of appliances [7], it is limited in two major aspects. Firstly, the detection of appliances with a power rating of less than 150 watts is unsupported ([7], [8]). Secondly, the approach is confined to appliances that emit repeatable transient patterns when their state is changed (referred to as *ON/OFF* or *state machine* appliances) [2], excluding variable loads from their detection.

Improvements to NIALM were achieved by collecting current and voltage samples at higher sampling rates [9]. This method is specifically suited for appliances characterized by significant overshoots in the power draw during their activation. Besides regarding only the real and reactive power components during transients, the approaches presented in [10] and [11] additionally regard harmonics that are collected during the occurrence of transients (i.e., during device startup or shutdown). Distinct features in the transient power waveforms have been found by isolating each harmonic and analyzing the spectral envelope over a fixed duration during the device startup. In addition to the analysis of current transients, the harmonic components of the steady-state current also serve as features for the classification of an attached appliance [12].

Because the aforementioned appliance identification algorithms have mostly been tailored to their application on data that has been collected at meter level, the approaches cannot resolve ambiguities when several identical appliances are present. As a result, metering equipment has been progressively moved closer towards the consumers in order to cater to both improved sensing resolution and less ambiguities. In a first step, power meters have been installed at circuit level, effectively overcoming the inability of single point sensing systems to monitor very small power consuming devices [13]. This increased granularity results in fewer appliances attached to each meter, hence a lower occurrence of indistinguishable devices can be expected. Furthermore, high-power devices, e.g., stoves or air conditioning units, are generally connected to dedicated circuits, and thus more sensitive sensors can be employed for monitoring devices with lower power draw. The deployment of metering equipment in individual wall outlets, as presented in [14], [15], is the next logical step towards the completely distributed deployment of sensing devices.

The final step in moving closer to the appliances to be classified is to attach sensor units to individual consumers. Despite the higher installation efforts, these distributed sensors can both be utilized to identify as well as control the connected appliance. The Plug [16], the ACme platform [17] and the Plugwise [5] are examples of such direct distributed sensing platforms, which monitor the power consumption of the plugged-in appliance. An identification of the attached loads is however unsupported by these platforms, despite the fact that they employ microcontrollers which can be programmed to locally process collected readings.

In this paper, we follow this latter approach of connecting metering devices between the wall outlets and the appliances to identify. In contrast to existing approaches, however, our implementation permits the extraction of a characteristic fingerprint from sensor data on a local scale and to relay this pre-processed fingerprint to a server for its further classification. Thus, our solution enriches distributed power metering solutions by local feature extraction methods which originate from the domain of NIALM.

III. Appliance Identification Architecture

Our architecture for the device classification is based on two major components, a distributed metering module and a processing server, and visualized in Fig. 1. The tasks of each component are discussed in the following subsections.

A. Distributed Metering

The general starting point to automatically discriminate between electrical devices, without making it necessary for the device to convey any information specifically, is to look at its electricity consumption behavior and extract distinctive features from it. While the voltage between the terminals of an appliance can be assumed to be of constant frequency and amplitude in the first approximation, the electric current flowing into a device has been shown to exhibit characteristics specific to the considered appliance. The capability to successfully identify a plugged-in electric appliance thus depends on the successful extraction of distinctive features from its current consumption, which we term the *fingerprint* of the appliance. Based on insights attained in related work (cf. [3], [4], [18], [19]) and our previous results [6], we focus on the *microscopic* current consumption in this paper, i.e., on characteristics of the current waveform in time scales of milliseconds.

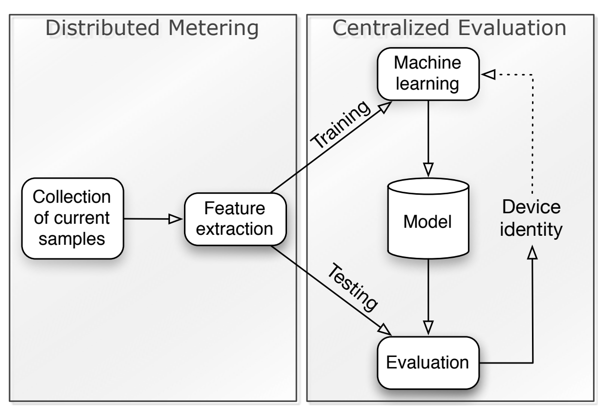
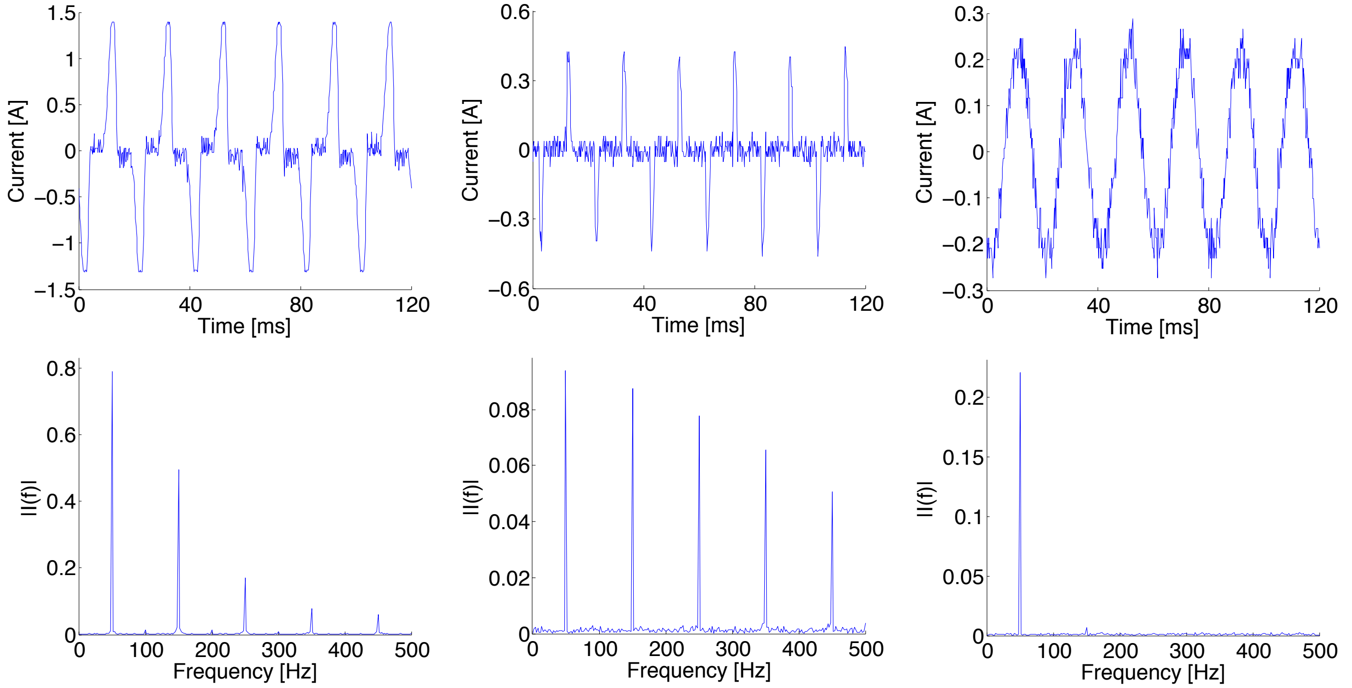


Fig. 1. Data collection and evaluation architecture



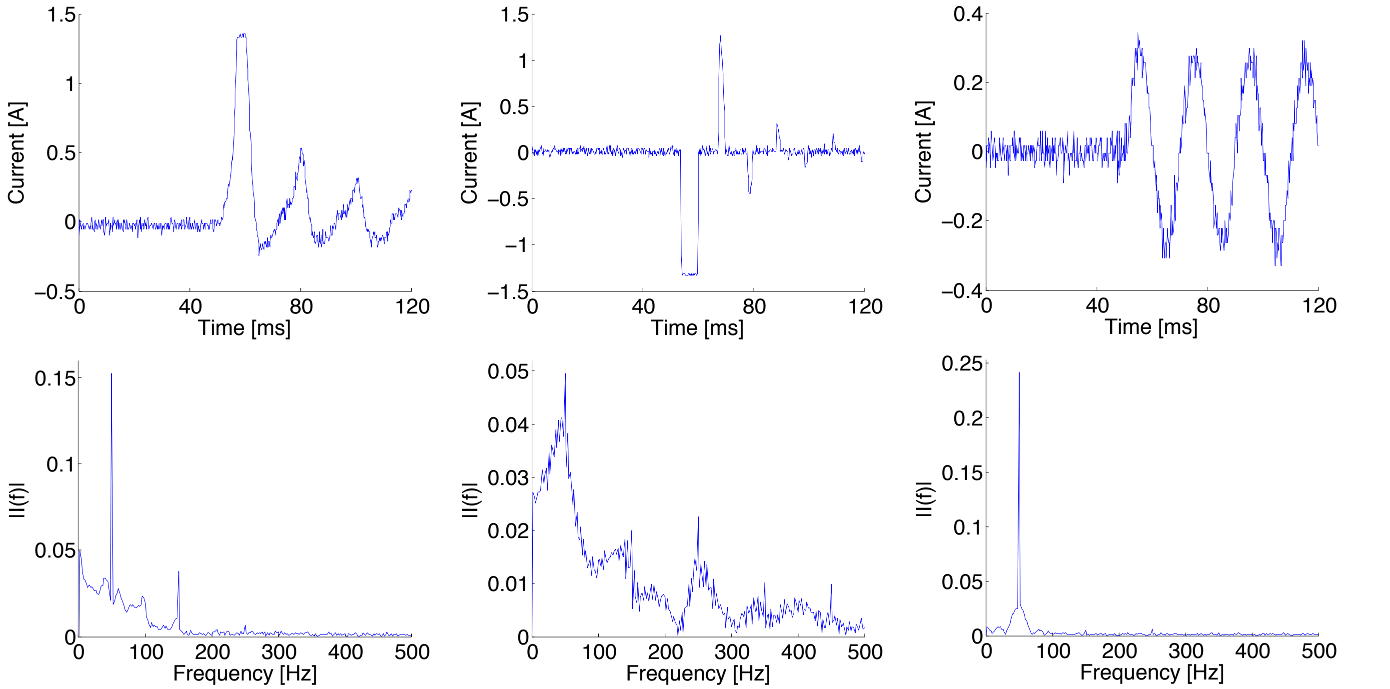
(a) Desktop computer (b) Fast Ethernet switch (c) Pedestal fan

Fig. 2. Steady-state current waveforms and corresponding frequency spectra of three electric appliances

For the collection of current samples, we use an extended version of our SmartMeter.KOM nodes. In contrast to the first revision of SmartMeter.KOM [6], which has been specifically designed to capture current flows at high sampling rates and allow for only minor local processing of the data, we have replaced the microcontroller to allow for more complex processing operations. More precisely, an Atmel microcontroller with 32kBytes of Flash memory, 2.56kBytes of RAM, and a clock frequency of 16MHz is now used. The nodes are placed into the mains connection of a number of appliances and form a wireless sensor network, through which they route all collected fingerprints to the sink node, which performs the evaluation step (cf. Sec. III-B). Identical to the original SmartMeter.KOM implementation, the current is captured at a sampling rate of 1.6kHz. More precisely, the embedded sensing system captures 512 current readings during a period of 320ms. In order to determine the phase shift between voltage and current, and thus enable the detection of capacitive and inductive loads, the SmartMeter.KOM nodes also annotate each current sample by whether it has been collected during the positive or negative half cycle of the full voltage sine wave.

In total, we calculate ten different features from each sampling window. Based on the current samples, we regard both the root mean square (*RMS*) and the arithmetic mean values. The latter is relevant because it enables the detection of appliances with asymmetrical current consumption during the voltage's half waves. The maximum absolute current value encountered in the sampling window as well as the phase shift between current and voltage for one single half wave are regarded as two further features. The remaining six features are extracted by applying a Fast Fourier Transform (*FFT*) over the sampling window. From its output, we use the magnitudes of the fundamental frequency and the first four odd harmonics (i.e., 150Hz through 450Hz for a mains frequency of 50Hz) as well as the DC component of the current waveform. In order to reduce the amount of data that needs to be transferred to the centralized evaluation server, the feature extraction is performed locally at the distributed metering unit. As a result of the local preprocessing, we reduce the traffic over the wireless channel from more than one kilobyte of raw sensor readings to a payload size of just 26 bytes.

Traces of the first six periods of the steady state current and the corresponding spectra are shown for three appliances in Fig. 2. Our decision to use the odd harmonics of the signal as distinctive features in our application is furthermore confirmed by the visual differences in the spectra depicted in the lower part of the figure. In addition to an analysis of steady-state currents, our system also regards the initial inrush current of appliances in order to support their classification. SmartMeter.KOM nodes comprise a solid state relay with zero-crossing detection, and can thus actuate an attached consumer in such a way that identical inrush currents can be observed repeatedly, only differing in terms of their sign (which has no impact on the FFT). We show three inrush current curves and the corresponding spectra in Fig. 3. While an increasing amount of noise can be observed in the spectra, the odd harmonics are still distinctive in the signal, and thus retain the applicability of our approach. The calculation of spectral components during the activation of an appliance is also motivated further by the spectrogram of the fast Ethernet switch shown in Fig. 4, where a visible difference between the inrush and steady state currents can be discerned.



(a) Halogen lamp (b) Fast Ethernet switch (c) Pedestal fan

Fig. 3. Inrush current waveforms and corresponding frequency spectra of three electric appliances

B. Centralized Evaluation Server

After the collection and preprocessing of the data, the collected fingerprints are forwarded to the evaluation server. The server uses data mining techniques in order to autonomously correlate collected fingerprints to the corresponding appliance type. The model is created using a supervised learning approach in which initially (i.e., during a *training* period) all incoming fingerprints are labeled with the type of the appliance from which they were captured. The server uses the Weka toolkit [20], which provides a large collection of classifiers and evaluation tools, and which we also use in order to assess the classification accuracies and the required time for the model building (cf. Sec. IV). The model retained at the server, which contains the correlation between a given appliance type, is extended with each received fingerprint and thus contains an autonomously established mapping between fingerprints and appliance types after the training phase.

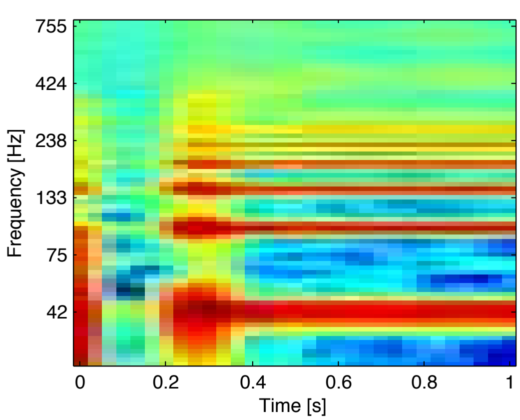


Fig. 4. Spectrogram of the fast Ethernet switch

Once the training phase has succeeded, the systems changes to the *testing* phase, in which incoming fingerprints are no longer annotated with the underlying appliance type. The sole operation of the server conducted during the testing phase is the matching of incoming fingerprints against the model and returning the corresponding appliance type. When the identity of an appliance is known, manifold notification and visualization options can be realized on the server, e.g., to notify the user about unnecessary energy consumption or to generate statistics of typical device operation schedules. It needs to be remarked that no further fingerprints are added to the model during the testing phase in our implementation.

The decision to collect all fingerprints and the model at a central instance has been made for several reasons. First and foremost, the computation and memory demand to create and dynamically extend the model poses resource requirements that exceed the capabilities of the SmartMeter.KOM nodes. The wireless transfer of fingerprints thus alleviates the need for expanding the computational power of each low-cost power sensing device by shifting the resource-intensive tasks to the server. Secondly, retaining a complete and up-to-date model at a central instance eliminates the need to distribute models to all devices whenever an update to the model has occurred at any node. Finally, the annotation of the collected traces with the device identities during the training phase can be done more conveniently on a user-accessible system with a graphical user interface.

IV. Evaluation

We evaluate our system in three consecutive analyses. First, we assess whether the harmonic features of the inrush or the steady state are more relevant to the classification accuracy, and determine the outcome when both are combined. In the second evaluation, we increase the number of devices for which traces have been collected, and evaluate the impact of the harmonic features on the classification accuracy. For the third and last experiment, we have collected 3,400 current waveforms using the developed system, and used them to analyze the classification accuracies and the classification duration of different implementations.

A. Impact of the Inrush Current

In this first analysis, we have collected 150 traces of load signatures from a set of six devices of five different types, as presented in Table I. The *computer monitor* class includes two different monitors (one Dell 20" and one Fujitsu Siemens 24" LCD screen) and hence comprises twice the number of samples. From each of the collected traces, the aforementioned harmonic content features were extracted for both the inrush and the steady state current, i.e., the current waveform captured five seconds after the device activation. The resulting training data thus comprises of the six FFT-based features, namely the magnitudes of the fundamental frequency, the odd harmonics, and the steady component. The major goal of this first evaluation is to compare the impact of the features from both the steady state and inrush currents on the classification accuracy. For the evaluation, we rely on the Bayesian network classifier (our analysis of the achievable classification accuracies is presented in Sec. IV-C).

The classification accuracy has been analyzed using a 25-fold cross validation of the training data. The resulting confusion matrices for inrush current traces (collected with the zero-crossing solid state relay), the steady state current waveforms, and their combination are compared in Table II. Confusion is observed when solely relying on the features of any of the two feature sets individually. These ambiguities could however be completely resolved when the combination of both current feature sets was regarded in our evaluation of the five device types. As a result of this first evaluation, the selected harmonic features have proven to be sufficient in order to classify this small set of appliances correctly. The absence of devices with similar nominal power consumptions in this evaluation however does not permit a generalization of the results, and motivates the following analysis.

B. Impact of the Harmonic (FFT-based) Features

In this second evaluation, we use a slightly larger set of device types and regard all of the ten features introduced in Sec. III-A. Based on the observations made in the first experiment, both the inrush and the steady state current waveforms have been regarded in this analysis. For our evaluation, we have collected the feature vectors of seven different household appliances, as listed in Table III. For each of the devices, 30 traces of their inrush and steady state currents have been collected each, such that an entire data set of 210 instances has been created.

In the evaluation, we assess the classification accuracy of three different machine learning algorithms, namely the J48, Naive Bayes, and Bayesian network algorithms. Furthermore, we investigate the effect of the additionally regarded features on the classification accuracy. Again, a 25-fold cross validation is used to analyze the data set. The confusion matrices for each of the algorithms as well as the overall classification accuracies for the given data set and the full feature vector set are presented in Table IV (J48), Table V (Naive Bayes), and Table VI (Bayesian network), respectively. Again, the Bayesian network shows the best classification accuracy of more than 98%.

In order to quantify the differences between relying solely on the six FFT-related features and the full feature set, we have

TABLE I
Appliance types and traces captured for the first evaluation

Device class	Number of traces
Fast Ethernet switch (A)	25
Desktop computer (B)	25
Computer monitor (C)	50
Pedestal fan (D)	25
Halogen desktop lamp (E)	25

TABLE II
Result of the device classification when the six harmonic features are being used to train a Bayesian network classifier

Input traces	Confusion matrix	Accuracy
Steady state	$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array}\begin{array}{c} \begin{array}{ccccc} \text{A} & \text{B} & \text{C} & \text{D} & \text{E} \end{array} \\ \begin{bmatrix} 24 & 0 & 1 & 0 & 0 \\ 0 & 25 & 0 & 0 & 0 \\ 0 & 0 & 30 & 19 & 1 \\ 0 & 0 & 3 & 20 & 2 \\ 0 & 0 & 2 & 2 & 21 \end{bmatrix} \end{array}$	80.0%
Inrush current	$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array}\begin{array}{c} \begin{array}{ccccc} \text{A} & \text{B} & \text{C} & \text{D} & \text{E} \end{array} \\ \begin{bmatrix} 25 & 0 & 0 & 0 & 0 \\ 0 & 24 & 0 & 0 & 1 \\ 0 & 0 & 50 & 0 & 0 \\ 0 & 0 & 0 & 25 & 0 \\ 0 & 0 & 0 & 0 & 25 \end{bmatrix} \end{array}$	99.3%
Both	$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array}\begin{array}{c} \begin{array}{ccccc} \text{A} & \text{B} & \text{C} & \text{D} & \text{E} \end{array} \\ \begin{bmatrix} 25 & 0 & 0 & 0 & 0 \\ 0 & 25 & 0 & 0 & 0 \\ 0 & 0 & 50 & 0 & 0 \\ 0 & 0 & 0 & 25 & 0 \\ 0 & 0 & 0 & 0 & 25 \end{bmatrix} \end{array}$	100.0%

TABLE III
Appliance types used in the second experiment

Device class	Nominal power draw
Incandescent light bulb (F)	60W
Compact fluorescent lamp (G)	11W
Laptop power supply (H)	368W (max.)
Computer monitor (I)	160W
Desktop fan (J)	40W
Audio amplifier (K)	100W
Immersion blender (L)	175W

TABLE IV
CONFUSION MATRIX, J48 CLASSIFIER

$$\begin{array}{cc} & \begin{array}{ccccccc} \mathrm{F} & \mathrm{G} & \mathrm{H} & \mathrm{I} & \mathrm{J} & \mathrm{K} & \mathrm{L} \end{array} \\ \begin{array}{c} F \\ G \\ H \\ I \\ J \\ K \\ L \end{array} & \begin{bmatrix} 30 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 30 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 28 & 2 & 0 & 0 & 0 \\ 1 & 0 & 6 & 23 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 30 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 30 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 27 \end{bmatrix} \end{array} \rightarrow \text{Accuracy } 94.3\%$$

TABLE V
CONFUSION MATRIX, NAIVE BAYES CLASSIFIER

$$\begin{array}{cc} & \begin{array}{ccccccc} \mathrm{F} & \mathrm{G} & \mathrm{H} & \mathrm{I} & \mathrm{J} & \mathrm{K} & \mathrm{L} \end{array} \\ \begin{array}{c} F \\ G \\ H \\ I \\ J \\ K \\ L \end{array} & \begin{bmatrix} 29 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 30 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 28 & 1 & 0 & 0 & 1 \\ 0 & 0 & 5 & 24 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 30 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 30 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 30 \end{bmatrix} \end{array} \rightarrow \text{Accuracy } 95.7\%$$

TABLE VI
CONFUSION MATRIX, BAYESIAN NETWORK CLASSIFIER

$$\begin{array}{cc} & \begin{array}{ccccccc} \mathrm{F} & \mathrm{G} & \mathrm{H} & \mathrm{I} & \mathrm{J} & \mathrm{K} & \mathrm{L} \end{array} \\ \begin{array}{c} F \\ G \\ H \\ I \\ J \\ K \\ L \end{array} & \begin{bmatrix} 30 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 30 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 28 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 28 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 30 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 30 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 30 \end{bmatrix} \end{array} \rightarrow \text{Accuracy } 98.1\%$$

conducted a supplementary analysis. The results are shown in Table VII and indicate that approximately the use of the full feature vector leads to approximately 10 percent better results as compared to an approach relying on harmonic content only. Besides confirming that our classification approach is not confined to data sets of five appliance types only, a major outcome of this second experiment is the fact that an educated selection of an appropriate classifier has significant impact on the classification accuracy. When using the full feature vector, confusion was only present for some of the traces collected from the laptop computer's power supply and the computer monitor.

C. *Classifier Selection*

In this third and last evaluation, we have established a larger input set by collecting 3,400 current waveforms from 16 different appliances. The goal of this evaluation is an assessment of the classification accuracy for different machine learning algorithms as well as an analysis of the time required to construct the model from the input data. The selected electric appliances used in this evaluation are listed in Table VIII. Apart from the LCD monitor (class C), a total number of 200 traces were collected for each appliance. To assess the classifier's behavior when devices of the same make and model are present in the input, two different LCD monitors of identical brand and model were traced 200 times each, resulting in 400 traces for the LCD monitor. The chosen devices were specifically selected such that devices with similar nominal power draw are present in the input data set, allowing us to assess the classification accuracy when the stated power demand indicates the possible presence of ambiguities.

In order to ensure comparability to real-world settings, where significantly more steady state operating fingerprints than inrush traces are likely to be collected, we only consider one device operation phase per appliance. In other words, more than 95% of the collected current waveforms were captured during the appliance's steady state operation. The analysis of the collected data is again done using a 25-fold cross validation of the complete training set. We have analyzed the classification accuracies and model construction times for nine well-known machine learning algorithms. The resulting values are tabulated in Table IX, which shows the figures of merit for each of the classifiers. With regard to the classification accuracy, the first interesting result is that all classifiers reach values above 99% for the given set of extracted features. In

TABLE VII
CLASSIFICATION ACCURACY FOR DIFFERENT FEATURE SETS

Classifier	Harmonics only	Full feature set
Bayesian Network	85.2%	94.3%
Naive Bayes	83.3%	98.1%
J48	84.3%	95.7%

TABLE VIII
APPLIANCE TYPES USED IN THE THIRD EXPERIMENT

Device class	Nominal power draw
Compact fluorescent lamp (A)	11W
Laptop power supply (B)	368W (max.)
LCD monitor (C)	160W
LCD TV (D)	150W
Audio amplifier (E)	100W
Immersion blender level 12 (F)	175W
Immersion blender level 01 (G)	35W
Bread slicer (H)	110W
Refrigerator (I)	90W
Freezer (J)	100W
CRT TV (K)	130W
LCD monitor (L)	230W
Fluorescent tube (M)	18W
Fluorescent tube (N)	40W
Halogen lamp (O)	42W
Laptop power supply (P)	390W (max.)

TABLE IX
CLASSIFICATION ACCURACY FOR DIFFERENT CLASSIFIERS

Algorithm	Cross validation accuracy (25 folds)	Model construction time
Bagging	99.88%	0.59s
Bayesian Network	**100%**	**0.20s**
J48	99.76%	0.17s
JRip	99.29%	0.51s
LogitBoost	99.29%	0.64s
Naive Bayes	99.97%	0.08s
Random Committee	99.94%	0.22s
Random Forest	99.97%	0.20s
Random Tree	99.79%	0.03s

TABLE X
RANKING OF THE EXTRACTED FEATURES BY INFORMATION GAIN

Rank	Feature
1	Average current
2	Magnitude of 50Hz harmonic
3	RMS value of the current
4	Peak current
5	Magnitude of 150Hz harmonic
6	Magnitude of 250Hz harmonic
7	Magnitude of 350Hz harmonic
8	Magnitude of 450Hz harmonic
9	Phase shift
10	Magnitude of DC offset

other words, the selected set of features appears to be highly suited for the correct classification in almost all cases. For reference, a ranking of the features' importance, based on their information gain, is provided in Table X. As the overall result, the Bayesian network classifier has the best result with no confusion between devices at all, and even its construction time lies within the top third of the probes. Concluding from these experimental observations, Bayesian networks are best suited for the problem at hand.

V. SUMMARY AND CONCLUSIONS

In many application domains, such as large-scale deployments (e.g., in office buildings) or sites with frequent changes to the connected equipment, the configuration process of distributed metering equipment becomes prohibitively time-consuming. As a result, self-configuration is a clear necessity for the successful and scalable realization of smart environments and/or demand side management. We support this autonomous configuration by presenting means to identify electric appliances based on their electric current consumption. Our distributed load monitoring system is based on embedded current monitoring devices, which collect current readings at a sampling rate of 1.6kHz and extract ten features from the collected data. A machine learning implementation of the server maintains a model which matches these fingerprints to the learned appliance types. The results of our evaluations show that a very high classification accuracy values of more than 98% can be achieved when the fingerprint of the inrush current is regarded in addition to the features extracted from the steady state current waveforms. Our evaluations have proven that we can even resolve ambiguities between devices with a similar power rating (e.g., a bulb and a fan of 60 watts each). With respect to the classifier, we have determined best results for the Bayesian network classifier, which showed the highest classification accuracy across all conducted evaluations.

ACKNOWLEDGEMENT

This research has been supported by the German Academic Exchange Service (DAAD) under the "Working Internship in Science and Engineering" program.

REFERENCES

[1] European Smart Metering Industry Group, "Smart Metering for Europe – A Key Technology to Achieve the 20-20-20 Targets," Online: http://www.esmig.eu/newsstor/smart-metering-for-europe-new-position-paper, 2009.

[2] G. W. Hart, "Residential Energy Monitoring and Computerized Surveillance via Utility Power Flows," *IEEE Technology and Society Magazine*, vol. 8, no. 2, 1989.

[3] 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.

[4] H. Chang, C. Lin, and J. Lee, "Load Identification in Nonintrusive Load Monitoring using Steady-State and Turn-on Transient Energy Algorithms," in *Proceedings of the 14th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 2010.

[5] Plugwise B.V., "Plugwise Circle," Online: http://www.plugwise.com/idplugtype-f/circle, 2010.

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

[7] G. W. Hart, "Nonintrusive Appliance Load Monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, 1992.

[8] 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.

[9] A. Albicki and A. Cole, "Algorithm for Non-intrusive Identification of Residential Appliances," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*, 1998.

[10] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong, "Power Signature Analysis," *IEEE Power and Energy Magazine*, vol. 1, no. 2, 2003.

[11] M. Berges, E. Goldman, H. S. Matthews, and L. Soibelman, "Enhancing Electricity Audits in Residential Buildings with Nonintrusive Load Monitoring," *Journal of Industrial Ecology*, vol. 5, no. 14, 2008.

[12] D. Srinivasan, W. Ng, and A. Liew, "Neural-network-based Signature Recognition for Harmonic Source Identification," *IEEE Transactions on Power Delivery*, vol. 21, no. 1, 2006.

[13] A. Marchiori and Q. Han, "Using Circuit-Level Power Measurements in Household Energy Management Systems," in *Proceedings of the 1st ACM Workshop on Embedded Sensing Systems For Energy-Efficiency In Buildings (BuildSys)*, 2009.

[14] T. Saitoh and Y. Aota, "Current Sensor-based Non-intrusive Appliance Recognition for Intelligent Outlet," in *Proceedings of the 23rd International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*, 2008.

[15] R. Morales, F. J. Badesa, N. García-Aracil, C. Perez-Vidal, and J. M. Sabater, "Distributed Smart Device for Monitoring, Control and Management of Electric Loads in Domotic Environments," *MDPI Sensors*, vol. 12, 2012.

[16] J. Lifton, M. Feldmeier, Y. Ono, C. Lewis, and J. A. 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)*, April 2007.

[17] X. Jiang, S. Dawson-Haggerty, P. Dutta, and D. Culler, "Design and Implementation of a High-Fidelity AC Metering Network," in *Proceedings of the 8th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), Track on Sensor Platforms, Tools, and Design Methods (SPOTS)*, 2009.

[18] L. K. Norford and S. B. Leeb, "Non-intrusive Electrical Load Monitoring in Commercial Buildings Based on Steady-state and Transient Load-detection Algorithms," *Energy and Buildings*, vol. 24, no. 1, 1996.

[19] S. N. Patel, T. Robertson, J. A. Kientz, M. S. Reynolds, and G. D. Abowd, "At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line," in *Proceedings of the 9th International Conference on Ubiquitous Computing (UbiComp)*, 2007.

[20] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA Data Mining Software: An Update," *SIGKDD Exploration Newsletter*, vol. 11, no. 1, 2009.