NoFaRe
Non-Intrusive Facility and Resource Monitoring

Matthias Kahl, Christoph Goebel, Anwar Ul Haq, Thomas Kriechbaumer
and Hans-Arno Jacobsen

4. D-A-CH Konferenz Energieinformatik
12.-13. November Karlsruhe

Department of Computer Science
Chair for Application and Middleware Systems (I13)
Motivation

What is the problem?

No information available about appliances and the electricity grid without intrusive measurements (e.g. smart plugs).

What is the goal?

A smart meter that enables appliance and electricity information retrieval with machine learning techniques.

What is the task?

Classification of appliances and power disaggregation.
What we would like to have

intrusive measurement

non-intrusive measurement
Use Cases

• Identification of currently running appliances
  Is this a toaster, TV or a hair dryer?

• Information retrieval from currently running appliances
  Power, on-duration, state patterns

• Detection of energy waste
  Out-of-behavior pattern

• Maintenance support
  Localizing and predicting of appliance faults

• Building safety
  Detection of unauthorized appliances usage
**State-of-the-Art Overview**

- Reinhardt et al. [1]: up to 98% accuracy (laboratory environment)
  - 3,400 high frequency samples
  - 16 appliances
  - 10 spectral features

- Jiang et al. [2]: around 90% accuracy (real and laboratory environment)
  - low amount of high frequency samples
  - 11 appliances with an
  - edge based startup transient recognition

- Yang et al. [3]: around 95% accuracy (laboratory environment)
  - 341 high frequency samples
  - 5 appliances
  - industrial loads

- Patel et al. [4]: around 85-90% accuracy (real environment)
  - ~3000 high frequency samples
  - ~19 appliances
  - household loads, up to 100kHz observations, only voltage transients
State-of-the-Art Discussion

- Too few data samples
- Too few appliances
- Often laboratory environment
- Mostly only household
- Non-public data sets
- Missing ground truth
- Low appliance type diversity
Our Contribution

- Online and real-time electricity and appliance information retrieval system
- Comprehensive approach to appliance identification, recognition and classification
- Preliminary evaluation on existing data sets
Appliance Classification & Identification

Data Acquisition
- non-intrusive 3-phases measurements

Preprocessing
- compensation calibration
- filtering

Feature Extraction
- power harmonics
- cos-phi

Training
- SVM, ANN

Classification
- training mode
- runtime mode
## Existing and Forthcoming Data Sets

<table>
<thead>
<tr>
<th>Household</th>
<th>Office</th>
<th>Industrial</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK-DALE</td>
<td>NoFaRe Office (to be collected)</td>
<td>NoFaRe Industrial (to be collected)</td>
</tr>
<tr>
<td>BLUED</td>
<td>NoFaRe Lab (to be collected)</td>
<td></td>
</tr>
<tr>
<td>REDD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAID</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- High sampling frequency (>10kHz)
- High sampling resolution (>12bit)
- Appliance event ground truth
Preprocessing

• Compensating phase shift
  Induced by measurement system

• Scaling of voltage & current

  current

  \[ I = I' \cdot \frac{N_{\text{turns}}}{R_{\text{burden}}} \]

  voltage

  \[ V = V' \cdot \frac{V_{\text{primary}}}{V_{\text{secondary}}} \cdot \frac{R_1 + R_2}{R_2} \]
Feature Extraction

short-time features

computed mid-time features

long-time features

statistical analysis
Learning & Classification

- Learning of feature states in *training mode*
- Use of machine learning (k-NN, SVM, ANN …)
- Classification of unknown appliance in *runtime mode*

<table>
<thead>
<tr>
<th>Class</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kettle1</td>
<td>0.13</td>
<td>0.86</td>
<td>0.34</td>
</tr>
<tr>
<td>Kettle2</td>
<td>0.12</td>
<td>0.91</td>
<td>0.35</td>
</tr>
<tr>
<td>Toaster1</td>
<td>0.33</td>
<td>0.31</td>
<td>0.75</td>
</tr>
<tr>
<td>Toaster2</td>
<td>0.34</td>
<td>0.28</td>
<td>0.68</td>
</tr>
<tr>
<td>?</td>
<td>0.14</td>
<td>0.88</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Experimental Setup

- Extracting start-up transients
- Feature extraction on start-up transients
- Classification of start-up transients

<table>
<thead>
<tr>
<th>Features</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.13</td>
<td>0.86</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>0.12</td>
<td>0.91</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>0.33</td>
<td>0.31</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>0.34</td>
<td>0.28</td>
<td>0.68</td>
</tr>
</tbody>
</table>
UK-DALE Data Set Results

- 37299 samples from
- 2 households
- 23 classes
- 7 discriminating features
- low diversity per class
- High amount of samples

Accuracy with Cross Validation
- with SVM: 93%
- with k-NN: 89%
UK-DALE Data Set Discussion

- Indirect startup ground truth
- Real-world scenario
- Sample counts vary highly across classes
- Low diversity per class

High accuracy because of low inner-class diversity and uneven class distributions
Plaid Plug Data Set Results

• 408 samples
• 11 classes
• 55 households
• 7 discriminating features
• High diversity per class
• Low amount of samples
• Accuracy with Cross Validation
  • with SVM: 92%
  • with k-NN: 88%
Plaid Plug Data Set Discussion

- Data inconsistently labeled
- Not a real-world scenario but isolated samples
- High amount of data unusable because of missing calibration information

High accuracy because of laboratory quality
Future Work & Discussion

- More datasets (REDD, …)
- Features and feature combinations
- Influence of different sampling rates on accuracy
- Different classifier (SVM, K-NN, ANN, HMM, …)
- Feature space transformation (PCA, LDA)
References


What we would like to have

- Many intrusive sensors for each appliance
- Simple measuring device
- High hardware effort
- Low Software effort

- One non-intrusive Sensors for all appliances
- Intelligent Measuring device
- Low Hardware effort
- High Software effort
Features

- Mean Power: 
  \[ P = \text{RMS}(I) \cdot \text{RMS}(V) - P_{pre} \]

- Signal to Signal-Mean Ratio: 
  \[ \text{mags} = \text{abs} \left( \text{fft}(I) \right) \quad \text{SSMR} = \frac{\text{max(mags)}}{\text{mean(mags)}} \]

- Inrush Current Ratio: 
  \[ \text{ICR} = \frac{I_{\text{startup}}}{I_{\text{stationary}}} \]

- Phase Shift: 
  \[ \text{PS} = \text{angle}(\text{mags}_{\text{current}}(50Hz)) - \text{angle}(\text{mags}_{\text{volt}}(50Hz)) \]
Features (2)

- Curve Shape Approximation:

- Harmonics:

- V-I Trajectory:
Outline

• Motivation
• State-of-the-Art
• Algorithmic approach
• Pattern recognition in NILM

• Experiments
• Results
• Future Work & Discussion
Feature Extraction

- Ultra short-time features
- Acquired short-time features
- Long-time features
- Statistical analysis