NONINTRUSIVE LOAD MONITORING:
Past, Present, and (hopefully) Future

Sean Barker, in collaboration with Sandeep Kalra, David Irwin, and Prashant Shenoy

University of Massachusetts Amherst
Energy Monitoring

- Energy monitoring important in smart buildings

- Where is energy used?
  - Where should smart homes focus their efforts?

- What can energy use tell us?
  - Occupancy – e.g., microwave running
  - Peak usage – e.g., multiple A/Cs coming on at same time
Implementing Energy Monitoring

- How do we implement energy monitoring?

- Approach 1: *monitor everything*
  - Deploy sensors on all devices
  - Collect data feeds in parallel

- Pros
  - Accurate (generally)
  - Conceptually straightforward

- Cons
  - Intrusive to homeowners
  - Prohibitive infrastructure overhead
  - Equipment, installation, maintenance
Nonintrusive Load Monitoring

- Approach 2: **nonintrusive load monitoring** (NILM)
- Monitor only house-level aggregate power
  - Decompose aggregate signal into individual loads using software

- **Pros**
  - Low overhead and cheap (single meter)
  - Infrastructure already being rolled out by utilities

- **Cons**
  - Nontrivial algorithmic challenge

How can we decompose the aggregate power signal?
Outline

- Motivation

- The classical approach: edge detection
  - [Hart92, Nonintrusive appliance load monitoring]

- A modern approach: hidden Markov models
  - [Kolter11, REDD]

- Current UMass research
  - Feature-based model detection

- Conclusions
Edge Detection

- Idea: characterize the aggregate trace as a series of [ON, OFF] events (‘edges’)
- Match ON and OFF events within the aggregate trace to determine when devices are active
Edge Detection (2)

- Include reactive power in edges
  - Distinguish devices with similar power usage

- ‘Signature Space’ of devices

---

**Fig. 6.** Normalized complex power signature space. Resistive appliances (water heater, iron, infrared light) appear on the real power axis. Motors have a reactive component. Apparently place a diode in series with the heating element to implement the low setting. This results in a half-wave rectified current waveform with a significant dc component. See [12] for a current waveform figure. However, we have encountered no appliance of significant interest to utilities that has a substantial dc current flow.
Supervised vs Unsupervised

- Do we know what devices are in the house?
  - **MS** (manual setup) - known devices
  - **AS** (auto setup) - unknown devices

- Most NILM approaches have been supervised

- Unsupervised NILM requires identifying unknown devices within the trace

\[ P_{\text{norm}} = \frac{120}{V}^2 p \]

\[ 1 \text{ Hz Normalized real, reactive power on each leg} \]

**Fig. 8.** Nonintrusive load monitoring (NILM) procedure.

- Measure Power and Voltage
- Normalize
- Edge Detection
- Cluster Analysis
- Build Appliance Models
- Track Behavior in Terms of Models
- Tabulate Statistics
- Appliance Naming

- Not required in MS-NALM
Generalizing Edge Detection

- Problem: edge-detection assumes two-state devices (ON and OFF)
  - Real-world devices have multiple power states

- Recent approaches model devices as Hidden Markov Models (HMMs) with a set of discrete states
### Factorial Hidden Markov Models (FHMM)

- **Factorial HMM**: HMM variant that conditions observations on set of independent Markov chains
- **Used to perform NILM in REDD work** [Kolter11]
- **Each load in the aggregate trace trained as HMM** (set of states and transition probabilities)
- **Disaggregation looks at aggregate trace as composition of the independent HMM models**

![Diagram of Factorial Hidden Markov Models](image-url)
REDD FHMM Results

- Total energy correctly assigned (second-to-second)

<table>
<thead>
<tr>
<th>House</th>
<th>FHMMM</th>
<th>Simple Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>Test</td>
</tr>
<tr>
<td>1</td>
<td>71.5%</td>
<td>46.6%</td>
</tr>
<tr>
<td>2</td>
<td>59.6%</td>
<td>50.8%</td>
</tr>
<tr>
<td>3</td>
<td>59.6%</td>
<td>33.3%</td>
</tr>
<tr>
<td>4</td>
<td>69.0%</td>
<td>52.0%</td>
</tr>
<tr>
<td>6</td>
<td>62.9%</td>
<td>55.7%</td>
</tr>
<tr>
<td>Total</td>
<td>64.5%</td>
<td>47.7%</td>
</tr>
</tbody>
</table>

- Two-week energy breakdown

![Pie charts showing energy contributions](chart.png)
Outline

- Motivation

- The classical approach: edge detection
  - [Hart92, Nonintrusive appliance load monitoring]

- A modern approach: hidden Markov models
  - [Kolter11, REDD]

- Current UMass research
  - Feature-based model detection

- Conclusions
NILM Issues: Model Complexity

- Overly simplistic load models
  - Edge detection ON-OFF models
  - Multiple states better, but how many states needed?

- Discrete states not always even applicable

- Want more accurate models of electrical loads
Scalability

- Number of loads in home
  - ~100 in a typical home
- Complex, noisy devices lead to complex, noisy signal

Real-world aggregate trace

‘Edges’ over 24 hours

Want robustness to noise and large numbers of loads
Online Operation

- Ideally, NILM should be ‘live’ (operate in real-time)
- HMM approaches poorly suited to online operation
- Expensive training and disaggregation
- Exponential number of hidden states

Want to operate inexpensively in close to real-time
Our Approach: Feature-based Load Models

- 1. Define a small set of features describing common attributes of real-world electrical loads

- 2. Design accurate models of loads in the home based on these features

- 3. Disaggregate individual loads within a trace by detecting feature-based models
Electrical Load Types

- Loads in home can be categorized by their electrical properties

- Resistive – heating elements
  - lights, toaster, oven

- Inductive – motors
  - fridge, freezer, A/C

- Non-linear – irregular current
  - computers, TVs

- Composite – series of other types
  - washing machine, dryer
Feature Types – Decays

- Load types display common features in their electrical usage

- Decays seen in many resistive loads
  - Initial peak usage followed by gradual decrease
Feature Types – Spikes & Cycles

- **Spikes** seen in many inductive loads
  - Very brief (~1s), high usage followed by immediate fall

- **Cycles** also seen in many inductive loads
Feature Types – Stable Min/Max

- Non-linear loads often show highly variable, but stable min/max (SMM) behavior
  - Rapid variations returning to a stable level

- Non-linear loads without a SMM often draw power within a random range
Feature Detection

- **Decays**: fit exponential decay functions against aggregate data (detect series of negative steps)
- **Spikes**: large positive steps followed by lesser (but still significant) negative steps
- **Cycles**: evenly-spaced spikes
- **Stable Min/Max**: dense clusters of positive or negative deviations from a stable level
- **Random Range**: dense clusters of oscillations (up/down step reversals)
Feature-based Disaggregation

- Design **device models** based on parameterized feature detectors
  - E.g., fridge with 20 minute cycles and 200 watt ON step following initial spike

- Identify individual devices within trace ordered by most identifiable devices
  - E.g., stable min/max devices tend to be highly visible
  - Devices with few features (e.g., lights) hardest to identify

- Device **extraction** rather than power decomposition (e.g., FHMM approach)
  - Can identify devices without decomposing entire trace
Feature-based Disaggregation Results

- Implemented our prototype **PowerPlay** and tested on data from our complete home deployment

- Modeled and disaggregated a set of representative devices over a day-long period

- Also disaggregated using FHMM approach
Total Energy Assigned

- Energy used over day-long period by individual devices

- Both PowerPlay and FHMM accurately assign energy
Second-level disaggregation accuracy from:
- 1. Individual device trace alone (self-matching)
- 2. Aggregate home trace (disaggregation)

- PowerPlay is more robust to a noisy aggregate trace
Disaggregated Trace Accuracy

- Freezer disaggregation (actual/PowerPlay/FHMM)

- Dryer disaggregation (actual/PowerPlay/FHMM)

- PowerPlay accurately reconstructs device traces
Performance

- PowerPlay operates in near real-time
  - ~2s to disaggregate one device over 24 hour period
  - Can be scaled to many devices without slowing down identification of existing devices
  - FHMM approach took >10x duration to disaggregate same 24 hour period
Conclusions

- Accurate, scalable, and real-time load disaggregation through feature-based device models

- Outperforms representative FHMM approach
  - More than 2x accuracy on reconstructed device traces
  - Faster, near real-time disaggregation

- Ongoing work modeling more devices and evaluating over longer time periods
Questions?

Sean Barker
sbarker@cs.umass.edu