A Dynamical Systems Approach to Energy Disaggregation

TRUST Autumn 2013 Conference

Roy Dong, Lillian Ratliff, Henrik Ohlsson, S. Shankar Sastry

UC Berkeley

October 9, 2013
Privacy-Aware Sampling Policies for Advanced Metering Infrastructures

- Analyzed the trade-off between sampling rate and control efficacy in demand response programs.
- Used sampling rate as a proxy for privacy.
Energy disaggregation, or non-intrusive load monitoring:

The desired signal is the energy signals for each individual device.
Energy disaggregation

Roy Dong, Lillian Ratliff, Henrik Ohlsson, S. Shankar Sastry
UC Berkeley

A Dynamical Systems Approach to Energy Disaggregation
Motivations

Benefits of disaggregated energy consumption data:

- Providing appliance-level feedback to energy consumers can achieve 20% energy savings in residential buildings, and sustain savings over the long-term.\(^1\)

- Improve efficacy of energy-saving programs and rebates via strategic marketing.\(^2\)

- Identify potential improvements in machine efficiency, e.g. repair or replacement, or operational efficiency, e.g. settings, use patterns.\(^3\)

---

\(^1\) Gardner and Stern 2008, Laitner et. al 2009.  
\(^2\) Armel et. al 2013.  
\(^3\) Hart 1992.
Motivations

Implications for privacy:

- How much device-level information is contained in the aggregate signal?
  - Affects who should have access to the data.
- A reasonable definition of privacy for the smart grid.
  - Assume the adversary has access to information about how devices work, and typical consumption patterns.
  - The adversary can only perform tractable algorithms on data.
Background

Approaches seen in previous work:

- Transient analysis.\(^4\)
- Hidden Markov models.\(^5\)
- Sparse coding.\(^6\)

Our contribution:

- We provide a novel framework for modeling the energy disaggregation problem.
- We phrase the energy disaggregation task as a filtering problem.
- We provide a recursive algorithm for energy disaggregation which is amenable to online deployment.
- We provide conditions under which we can recover the best solution.

---

\(^6\) Kolter and Ng 2010.
Problem statement

We are given the aggregated power consumption signal: $y[t]$ for $t = 0, 1, \ldots, T$

$$y[t] = \sum_{i=1}^{D} y_i[t] \text{ for } t = 0, 1, \ldots, T$$

- $D$ is the number of devices in the building.
- $y_i$ is the power consumption signal of device $i$.

From $y$ and some prior knowledge, we wish to recover $y_i$. 

Roy Dong, Lillian Ratliff, Henrik Ohlsson, S. Shankar Sastry 
UC Berkeley
A Dynamical Systems Approach to Energy Disaggregation
With a prior on $u$, our goal is to find the best estimate of the input given our observed output.
Implementing this framework requires two phases:

- **Training phase:** Acquire models and prior.
- **Energy disaggregation phase:** Find best inputs given models and prior.
Training phase

- Given training data $z_i$.
- Use heuristic methods to estimate $u_i^z$.
  - If we assume the inputs are piecewise constant, we can use change detection methods.
- Use system identification results to find a model $h_i$, such that $z_i = h_i(u_i^z)$.
- Construct prior based on empirical distribution.
Energy disaggregation phase

Given:

- System models: $h_i$ for $i = 1, 2, \ldots, D$
- Prior on our inputs.

Assumptions

- The devices in our building are a subset of the devices $\{1, 2, \ldots, D\}$.
- Our models accurately capture device dynamics.
Energy disaggregation problem

\[ \arg \min_{\hat{y}, u} \quad L(\hat{y}, y_m) + g(u) \]
subject to \[ \hat{y}_i = h_i(u_i) \]
for \( i \in \{1, \ldots, D\} \)
\[ \hat{y} = \sum_{i=1}^{D} \hat{y}_i \]

- \( y_m \) is the measured power consumption.
- \( \hat{y} \) is the estimated power consumption.
- \( L \) is a loss function penalizing deviations of \( \hat{y} \) from \( y_m \).
- \( g \) is a regularization on the input that incorporates our priors.

Note:
- \( h_i \) reflects the dynamics of the devices.
- \( g \) reflects our priors on the usage patterns of devices.
Energy disaggregation phase

Definitions

A **segment** is an interval in which $\mathbf{u}$ is constant. $\delta \in \{0, 1\}^T$ is a **segmentation** of $\mathbf{u}$ if:

$$\delta[t] = \begin{cases} 1 & \text{if } \mathbf{u}[t] \neq \mathbf{u}[t - 1] \\ 0 & \text{otherwise} \end{cases}$$
Energy disaggregation phase

- In our problem, $\delta$ is unknown.
- We break up the problem into a segmentation problem and a simpler optimal control problem.
  - This is a hybrid optimal control problem that has been studied; we use results from the adaptive filtering literature.
We maintain a filter bank $\mathcal{F}$ to solve this problem.

- Each filter $f \in \mathcal{F}$ corresponds to a segmentation $\delta_f$.
- For each filter, we optimize across possible inputs given a segmentation.

Intractable: maintain an $f \in \mathcal{F}$ for each possible segmentation $\delta_f$.
Relaxation: consider a subset of possible segmentations.
A segmentation $\delta$ can be thought of as a leaf node on a binary tree of depth $T$. (This lends itself nicely to a recursive algorithm which can be run online!)
Relaxations: Don’t expand every path and prune off branches that are not promising.
Under some reasonable assumptions on the form of our prior, the accuracy of our models, and the minimal length of a segment, we have theoretical results:

**Theorem:** Optimality of the proposed algorithm’s branching policy.

Selectively branching only the optimal estimates at all times $t < T$ will guarantee any optimal estimate will be in the filter bank at time $T$.

**Corollary:** Optimality of the proposed algorithm’s pruning policy.

If an optimal estimate at time $T$ is has sufficiently high likelihood for all time $t < T$, then it will be in the filter bank at time $T$. 
Experiment

Small-scale experiment:

Sampling rate: 12 Hz.
The left is measured data, the right is estimated from our algorithm.
In summary:

- We use a dynamical systems approach to the energy disaggregation, allowing us to phrase the task as a filtering problem.
- We provide a tractable, recursive algorithm for energy disaggregation which is amenable to online deployment.
- Using ideas from adaptive filtering, we can relax the original exponential problem while guaranteeing we still recover the exact solution.

Future work:

- Application to a larger scale dataset.
- Use of models and priors in other contexts, e.g. demand response programs.
- Develop a metric for how ‘good’ a disaggregation algorithm is.