Inferring the behavior of distributed flexible electric loads

Greg Ledva, Laura Balzano, & Johanna Mathieu
Department of Electrical Engineering & Computer Science
University of Michigan

Funded by NSF (ECCS-1508943)
Load Served by a Substation

![Graph showing load served by a substation over time with amplitude in MW on the y-axis and time of day on the x-axis.]

5/12/16  J. Mathieu, University of Michigan
Load Served by a Substation

Time of Day

Amplitude [MW]

Total
Air Conditioning
Other
In this talk, we use measurements of real power only. We could consider additional measurements (reactive power, voltage, etc.) from multiple meters at different points in the distribution network.
Why disaggregate the substation load?

• Load coordination feedback

**Bi-directional communication**

**Broadcast control, aggregate measurements**
Why disaggregate the substation load?

• Load coordination feedback
  – (noisy) measurements of the aggregate power of coordinated loads are assumed in Mathieu et al. 2013; Can Kara et al. 2013; Bušić and Meyn 2016; Callaway 2009; ...

![Diagram](image)
Why disaggregate the substation load?

Additional uses in demand response...
• Load aggregator bidding
• Demand response event signaling (when/how much)

Beyond demand response...
• Energy efficiency via conservation voltage reduction
  – Disaggregate by load type
• Contingency planning
  – Disaggregate motor loads
• Reserve planning
  – Disaggregate PV production
Connections to other problems


**Problem:** Infer individual load behavior from a single power measurement (usually) sampled at high frequency (10kHz-1MHz) from the household main

**Solution approaches:** offline algorithms including change detection, supervised learning, unsupervised learning
Key differences

• We assume measurements at the substation, not the household

• We infer aggregate load (e.g., all air conditioning load), not individual load behavior

• We solve the problem online, not offline

• We use lower frequency measurements (e.g., taken every second to minute)

• In some cases, we may get to be “intrusive,” but not in this talk!
Possible Methods

• State estimation
  – Linear techniques require linear system models
  – Nonlinear techniques can be computationally demanding

• Online learning
  – Data-driven, model-free

• Hybrid approach: Dynamic Mirror Descent
  [Hall & Willet 2015]
  – Admits dynamic models of arbitrary forms
  – Optimization-based method to choose a weighted combination of the estimates of a collection of models
Outline

• Dynamic Mirror Descent
• Problem setting: Plant data/models
• Algorithm Models
• Results
• Next steps
Dynamic Mirror Descent

• Mirror Descent: online algorithm to estimate a fixed state


1. Compute the error between the model predictions and the measured data (i.e., loss function $\ell_t(\hat{\theta}_t^i, y_t)$)

2. Update the state in the direction of the negative gradient of the loss function

$$\tilde{\theta}_t^i = \arg \min_{\theta \in \Theta} \eta_t \left( \nabla \ell_t(\hat{\theta}_t^i, y_t), \theta \right) + D \left( \theta \| \hat{\theta}_t^i \right)$$
3. Use the estimated states to evaluate the models for the next time step
\[
\hat{\theta}_t^{i+1} = \Phi_t^i(\hat{\theta}_t^i)
\]

4. Compute a weighted version of the estimates
\[
\hat{\theta}_t + 1 = \sum_{i=1}^{N^{mdl}} w_{t+1}^i \hat{\theta}_{t+1}^i.
\]

5. Update the model weights
\[
\omega_{t+1}^i = \frac{\lambda}{N^{mdl}} + (1 - \lambda) \cdot \frac{w_t^i \exp \left( -\eta^r \ell_t \left( \hat{\theta}_t^i, y_t \right) \right)}{\sum_{j=1}^{N^{mdl}} w_t^j \exp \left( -\eta^r \ell_t \left( \hat{\theta}_t^j, y_t \right) \right)}
\]
Algorithmic guarantees

- **Regret**: performance with respect to a comparator $\theta_T$

$$R_T(\theta_T) \triangleq \sum_{t=1}^{T} \ell_t(\hat{\theta}_t) - \sum_{t=1}^{T} \ell_t(\theta_t).$$

- Often the comparator is the performance of a batch algorithm

- Hall and Willet derive bounds on the regret and show that for many classes of comparators regret scales sublinearly in $T$
**Problem Setting: Plant Data/Models**

- **Air conditioners:** 1000 equivalent thermal parameters (ETP) models, i.e., three-state hybrid models [Sonderegger 1978]

\[
\theta^i_{t+1} = A^i \theta^i_t + B^i m^i_t + E^i d^i_t
\]

\[
m^i_{t+1} = \begin{cases} 
0 & \text{if } \theta^a_{t+1} < \theta^{set,i} - \theta^{db,i} / 2 \\
1 & \text{if } \theta^a_{t+1} > \theta^{set,i} + \theta^{db,i} / 2 \\
m^i_t & \text{otherwise}
\end{cases}
\]

\[
P^i_t = (|Q^{h,i}| m^i_t) / \eta^i
\]

where \( \theta^i_t = [\theta^a_{t,i} \, \theta^m_{t,i}]^T \)

- **Other loads:** data from Pecan Street Inc. Dataport
Problem Setting: Plant Data/Models

![Graph showing time of day vs. amplitude (MW)]

- **Total**
- **Air Conditioning**
- **Other**

Time of Day:
- 12AM
- 3AM
- 6AM
- 9AM
- 12PM
- 3PM
- 6PM
- 9PM
- 12AM

Amplitude [MW]:
- 0
- 5
- 10
- 15
- 20

5/12/16

J. Mathieu, University of Michigan
Algorithm Models

→ 57 models
→ $\hat{y}_t = \hat{y}^c_t + \hat{y}^{uc}_t$
  ACs + Other
→ AC models initialized at actual value, and run open-loop

**Model Set Estimates**

<table>
<thead>
<tr>
<th>Time of Day</th>
<th>Power (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12AM</td>
<td>20</td>
</tr>
<tr>
<td>3AM</td>
<td>19</td>
</tr>
<tr>
<td>6AM</td>
<td>18</td>
</tr>
<tr>
<td>9AM</td>
<td>17</td>
</tr>
<tr>
<td>12PM</td>
<td>16</td>
</tr>
<tr>
<td>3PM</td>
<td>15</td>
</tr>
<tr>
<td>6PM</td>
<td>14</td>
</tr>
<tr>
<td>9PM</td>
<td>13</td>
</tr>
<tr>
<td>12AM</td>
<td>20</td>
</tr>
</tbody>
</table>

**TABLE I**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient of Performance</td>
<td>[-3]</td>
</tr>
<tr>
<td>Internal Air Heat Gain</td>
<td>[kW]</td>
</tr>
<tr>
<td>Air Heat Capacitance</td>
<td>[C]</td>
</tr>
<tr>
<td>Envelope Conductance</td>
<td>[C]</td>
</tr>
<tr>
<td>Temperature Dead-band</td>
<td>[°C]</td>
</tr>
<tr>
<td>Temperature Set-Point</td>
<td>[°C]</td>
</tr>
</tbody>
</table>

5/12/16
J. Mathieu, University of Michigan
1000 two-state hybrid models
[Chong & Debs 1979; Ihara & Schweppe 1981]

\[ \theta_{t+1}^i = A^i \theta_t^i + B^i m_t^i + E^i d_t^i \]

\[ m_{t+1}^i = \begin{cases} 
0 & \text{if } \theta_{t+1}^a,i < \theta_{\text{set},i} - \theta_{db,i}/2 \\
1 & \text{if } \theta_{t+1}^a,i > \theta_{\text{set},i} + \theta_{db,i}/2 \\
m_t^i & \text{otherwise}
\end{cases} \]

\[ P_t^i = (|Q^{h,i}| m_t^i)/\eta^i \]

where \[ \theta_t^i = \theta_t^{a,i} \]
Algorithm Models: Air conditioners - 1

![Graph showing Amplitude [MW] vs. Time of Day with curves labeled $y_t$ and $\tilde{y}^c_{t,Hyb}$]
Algorithm Models: Air conditioners - 2

Linear Time Invariant (LTI) aggregate system models [Mathieu et al. 2013]

\[ x^{i}_{t+1} = A^{i} x^{i}_{t} \]  \[ \hat{y}^{c,\text{LTI},i}_{t} = C^{i} x^{i}_{t} \]

\[ i \in \mathbb{N}^{\text{temps}} \]

\[ \hat{y}^{c,\text{LTI},i}_{t} = C^{i} x^{i}_{t} \]  \[ i \in \mathbb{N}^{\text{temps}} \].
Algorithm Models:  
Air conditioners - 3

Linear Time Varying (LTV) aggregate system models
[Mathieu et al. 2015]

\[ x_{t+1} = A_t \ x_t \]
\[ \hat{y}_{t}^{c,LTV} = C_t \ x_t. \]

![Graph showing temperature variations throughout the day with LTV models](image-url)
Algorithm Models: Air conditioners - 3

Time of Day

Amplitude [MW]

0 12AM 6AM 12PM 6PM 12AM

0 1 2 3 4 5 6

$\hat{y}_t^c, LTV1$  $\hat{y}_t^c, LTV2$  $y_t^c$
Algorithm Models: Other loads

Time of Day

Amplitude [MW]

$y_t^{uc}$ $\hat{y}_t^{uc,Mon}$ $\hat{y}_t^{uc,Tues}$ $\hat{y}_t^{uc,Wed}$

rescale

5/12/16
Prediction Results

Amplitude [MW]

Time of Day

12AM 6AM 12PM 6PM 12AM

0 2 4 6

5/12/16

Weightings:
each color is a different model
Prediction Results: Better Models

Amplitude [MW]

Time of Day

12AM  3AM  6AM  9AM  12PM  3PM  6PM  9PM  12AM

0  2  4  6

\( y_t \)

\( \hat{y}_t \)
Weightings: Better Models

![Graph showing weightings for different models over time of day.](image)
Prediction Results: Bad Models

- All “other load models” are too low.
## Results: Summary

<table>
<thead>
<tr>
<th>Case</th>
<th>RMS Error (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benchmark:</strong> Use current outdoor temperature, LTI models, and interpolation to predict</td>
<td>738</td>
</tr>
<tr>
<td><strong>DMD Case 1:</strong> Includes every combination of aggregate air conditioner model and “other load model”</td>
<td>264</td>
</tr>
<tr>
<td><strong>DMD Case 2:</strong> Case 1 models plus a smoothed version of the actual “other loads”</td>
<td>211</td>
</tr>
<tr>
<td><strong>DMD Case 3:</strong> Case 2 models plus more accurate models of the aggregate air conditioning load over time periods where the other models are less accurate</td>
<td>175</td>
</tr>
<tr>
<td><strong>DMD Case 4:</strong> Includes “other load models” that underestimate the “other load”</td>
<td>1392</td>
</tr>
</tbody>
</table>
Results: Varying
Algorithm Parameters

Recall: \[ w^{i}_{t+1} = \frac{\lambda}{N_{mdl}} + (1 - \lambda) \cdot \frac{w^{i}_{t} \exp \left(-\eta^{r} \ell_{t} \left(\theta^{i}_{t}, y_{t}\right)\right)}{\sum_{j=1}^{N_{mdl}} w^{j}_{t} \exp \left(-\eta^{r} \ell_{t} \left(\theta^{j}_{t}, y_{t}\right)\right)} \]

![Graph showing RMS error vs \( \eta^{r} \)]

- Under-fitting
- Over-fitting
Next steps

- Investigate more realistic settings
- Develop better load models
- Improve the algorithm, e.g., alternative weighting functions
- Investigate identifiability
- Incorporate additional measurements (reactive power, voltage) into the approach
Conclusions

- Dynamic Mirror Descent (DMD) enables us to solve the substation disaggregation problem leveraging dynamical models of arbitrary form.

- DMD can work well (on simple examples); however, it is easy to find instances where it does not work well.