Energy Consumption Lifestyle Segmentation Using Hourly Electricity Data

Jungsuk Kwac, June Flora and Ram Rajagopal

Stanford Sustainable Systems Lab
Electrical Engineering
Civil and Environmental Engineering
Stanford University
Outline

• Introduction
• Methodology
• Segmentation results
• Conclusion and future work
Previous customer segmentation

Psychographic segmentation based on survey

**Issues:**
Rare actual energy usage info,
Low energy program yield (10-20%)
New opportunity with AMI

Advanced metering infrastructure deployment

New opportunity:
Interval energy consumption data (1 hour, 15 min)

New features available
ex) Peak hour, usage pattern, base consumption

http://www.eia.gov
1. Segment the customers with relevant features

⇒ *The features should represent energy consumption lifestyles well.*

2. The methodology should be
   a. *Innovative*
   b. *Scalable*
   c. *Analytic*
• Introduction
• Methodology
• Segmentation results
• Conclusion and future work
Segmentation flow

Raw Smart Meter Data → Pre-processed Dictionary → Encoding System → Feature Extraction → Clustering

Smart meter data

Dictionary

Hour

kWh

0.0 1.5

Hour

5 10 15 20
Load shape dictionary generation

- Load shape dictionary generation flow

Smart meter data

Normalize by daily consumption

Adaptive K-means + Hierarchical clustering
• Introduction
• Methodology
• Segmentation results
• Conclusion and future work
• **Anonymized smart meter data** provided by Pacific Gas and Electric Company (PG&E)

• Climate zones 1, 2, 3, 4, 11, 12, 13 and 16

• 123,150 households of 1 year data:
  45MM daily profiles

http://www.energy.ca.gov
Avg daily consumption
~ the size of household

Segmentation by usage volume

Probability density function of avg daily consumption

#958: 8.79%
#869: 7.09%
#742: 5.92%
#644: 5.54%
#891: 2.83%
#737: 2.72%

#715: 3.98%
#918: 2.94%
#811: 2.16%
#785: 1.96%
#844: 1.79%
#832: 1.67%
Segmentation by variability

**Entropy of load shape codes**
~ consumption pattern variability
### Load shape segment ~ peak hours

<table>
<thead>
<tr>
<th>Load shape segment</th>
<th>Morning peak (M: 4:00-10:00)</th>
<th>Daytime peak (D: 10:00-16:00)</th>
<th>Evening peak (E: 16:00-22:00)</th>
<th>Night peak (N: 0:00-4:00 and 22:00-24:00)</th>
<th>Dual peak Morning &amp; Evening (DU M&amp;E)</th>
<th>Dual peak Evening &amp; Night (DU E&amp;N)</th>
<th>Dual peak Daytime &amp; Evening (DU D&amp;E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg daily consumption (kWh)</td>
<td>20.73</td>
<td>24.06</td>
<td>19.73</td>
<td>10.78</td>
<td>15.81</td>
<td>11.23</td>
<td>24.49</td>
</tr>
<tr>
<td>Proportion households</td>
<td>4.29%</td>
<td>3.55%</td>
<td>72.98%</td>
<td>9.89%</td>
<td>9.24%</td>
<td>0.01%</td>
<td>0.04%</td>
</tr>
</tbody>
</table>
Segmentation by multiple features

Multiple features
Ex) usage volume and variability

<table>
<thead>
<tr>
<th></th>
<th>Stable</th>
<th>Moderate</th>
<th>Light</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy</td>
<td>13.03%</td>
<td>11.62%</td>
<td>0.47%</td>
<td>25.11%</td>
</tr>
<tr>
<td>Moderate</td>
<td>9.42%</td>
<td>30.46%</td>
<td>10.05%</td>
<td>49.93%</td>
</tr>
<tr>
<td>Light</td>
<td>2.51%</td>
<td>7.85%</td>
<td>14.60%</td>
<td>24.96%</td>
</tr>
</tbody>
</table>
• Introduction
• Methodology
• Segmentation results
• Conclusion and future work
• Check the goal again

Provided segmentation examples with relevant features.

The methodology is innovative, scalable, analytic.
Future works

- Try **other features** that can represent household lifestyles

- Various applications
  - ex) load shape prediction, load prediction

- With feedback data, test the implications
Thanks!!!