Load Shape Clustering of Residential Smart Meter Data

Ling Jin, Sam Borgeson, Dan Fredman, Liesel Hans, Sid Patel, Anna Spurlock, Annika Todd
October 2016

Behavior Analytics: combine behavioral theories with cutting-edge data science
Interpret meter data as product of preferences and behaviors
- How should we group household energy behaviors?
- When do households use most of their energy? Peaks?
- How diverse are customer load shapes?
- Etc.

Project Overview: Data
- One utility, residential customers
- Hourly smart meter data
- 100,000 households, 2.5 years
- Some survey data

Our solution: Behavior Analytics
- Identify a finite set (aka dictionary) of representative load shapes that best describe all observed shapes
- Interpret these representative shapes in terms of scheduling, occupancy, equipment ownership, and patterns of human behavior
A load shape is the pattern created by 24 hours of demand data.
Shapes are produced by patterns of occupancy, equipment ownership, and behavior.

Clustering approach

Normalization (load shapes)
Adaptive K-means
Hierarchical Merging

De-min

Hourly smart meter data ("customer-day")

Dictionary load shapes

Suitable parameters:
- Error threshold
- Violation rate

Truncation

Top 12 Clusters

Cluster coverage

Annual membership count (%) by each cluster
Cumulative load shape coverage (%)
Can we create customer groups?

Distribution of Shapes per Customer

Energy “personalities”

Breakdown of shape characteristics

Shape-Groups: Percent of all Customer-Days

Peak timing and magnitude

Energy "personalities"
Seasonal load shape diversity

Temperature driven load shape diversity

Contact:
Sam Borgeson
sam@convergenceda.com

Ling Jin
ljin@lbl.gov
Clustering Results

- Applied clustering to pre-treatment year:
  - \(\sim 10^5\) customers, \(\sim 3 \times 10^7\) load shapes (customer-days)
  - 99 representative clusters/shapes
  - 30% violation rate to error threshold 0.3.

% of shapes by count for each representative clusters

Representative clusters sorted by counts

99 Load Shape Clusters
Top 6 shapes by season

Weekday vs. Weekend

Normalization Flattens Load Shapes
99 Load Shapes: Now What?

- Distill 99 into more usable groupings
  → How should they be grouped?
- Behavioral dimensions in TOU context:
  - Number of peaks
  - Timing of peaks
  - Activity level during TOU period
- Can we blend qualitative intuition with data-driven processes?

1. Find breaks in data
2. Select high activity periods
3. Find local maxima
4. Group load shapes into descriptive categories
Can we create customer groups?

Examples:
• “Consistent” HHs = customers with only 1 or 2 shapes over all summer, non-holiday weekdays
  • Typically HHs with only 1 or 2 adults in the home
• What increases variation in a customer’s number of shape-groups?
  • Kids
  • Electric Dryer

Next Steps
• How to assign a household to a group?
  • How different are the groups?
  • How consistent (i.e., reliable) are HHs?
    => Further consolidation of shapes
• Are there any useful patterns via survey responses?
• How much does weather matter?
  • What is the distribution of shapes on event-like days?
• Are there patterns by day of week?