

Clustering Residential Customers by their Load Profiles for Demand Response Events

Pattanut Chanpiwat^a, Steven A. Gabriel^{a,c}, Michael Siemann^b, Rachel Moglen^a

^aUniversity of Maryland, College Park, Maryland

^bWhisker Labs, Germantown, Maryland

^cNorwegian University of Science and Technology, Trondheim, Norway,
Energy Transition Programme, <https://www.ntnu.edu/energytransition>

The study used kmeans and hierarchical clustering analyses to cluster residential households into useful homogeneous groups based on time-series 24-hour energy use and HVAC load. The understanding of customer classification and having electricity load profiles for different customer classes are top priorities for retail energy providers (REPs), so they can plan and conduct demand response (DR) effectively and efficiently.

Most clustering techniques require some prior assumptions such as the number of subgroups or density of subgroups. To eliminate the need for such assumptions, the clustering algorithms used in this research use sensitivity analysis to vary the number of customer clusters from 2 to 15 based on a set of 24 validation indices. The research is developed and validated using a dataset of roughly 14,000 residential households in 2017 located in Houston, Texas, USA.

Kmeans and Hierarchical cluster analysis are unsupervised techniques that there currently is no empirical means to validate number of cluster segment. The number of cluster segment derived either from expert practical standpoint or purely mathematical standpoint. In an expert-involved case, there is a need for psychological validation to get the most compelling outcome. That is context-dependent. The results is practically meaningful to the real world. In the mathematical case, the practice is to vary the numbers of cluster segment from 2 up to the maximum number of cluster segment, which is 15 in this research. Then, the result of each iteration is validated against as many measuring indices as possible. Out of the total indices, the majority rule helps selecting the final number of cluster segment. Each index is so sensitive to different types of input dataset. In this study, we validated results on different combinations of 9 methods, 5 distance measure, and 24 validation indices. The study simulated 500 runs of 1,000 randomly selected residential customers to be classified by 9 cluster analysis method and 5 distance measure. In each of those cases, it validated against 24 indices

The residential customers classification and load profiles obtained from this study are important for decision support to conduct demand response on segmented homogeneous group of customers. The study found that there are 6 distinct clustering classes that REPs can base on to create different load profiles. The right classification helps REPs to conduct demand response on customers differently according to their classes. In return, the REP is better to mitigate risk of unsuccessful DR, reduce customers opt-out, and increase saving both on physical energy demand and electrical bills for customers.