Pattern recognition and segmentation of smart meter data

Background

Electricity smart meters provide half hourly consumption data for residential and business consumers. This data potentially brings great possibilities for marketing or forecasting electricity demand, but also brings challenges in terms of the volumes of data.

In order to work with manageable data sets, electricity retailers aim to group customers into segments that are alike in some way, and aggregate the consumption data for that segment.

Choosing meaningful segments for a particular purpose could be done on the basis of:

» Categorical data such as whether the property is residential or commercial, and where it is located

» Load characteristics such as total annual load or ratio of peak load to average load

» Half-hourly load data, giving the ‘shape’ of a particular customer

This project is aimed at the final category, which is the least well understood. It is proposed that reference profiles are defined, and then the load data for individual customers is then compared to that reference profile, resulting in a score indicating how well the customer’s profile fits the reference profile.

The reference profile might be the average of all the customers already in that segment for example, or an archetypal customer profile.

Illustrative profile comparison

![Image of load profile comparison](image-url)
Research questions

Primary questions

» How can we define a reference profile from the load data for a population? Is there a better way than to take the average?

» What simple measures of distance are effective to give a score for the ‘closeness of fit’ between a customer profile and a reference profile? An effective measure might correspond well with what a human would see to be a good match.

Some sub-questions are:

» How to cope with differences of scale between the reference profile and customer profile and get to the ‘essence’ of pattern-matching

» How to cope with idiosyncrasies in customer data, for example holidays, air-conditioner breakdowns etc

Secondary questions

Some other interesting areas for research:

» A customer’s consumption consists as the sum of usage from a number of electrical. Is it possible to decompose customer profiles into their usage from these appliances? Generally historical data will not exist at this level.

» Given that weather is the key driver of electricity demand, how might weather data be used as part of this process?