Feature extraction to characterise and cluster the energy demand of UK retail premises

Ramon Granell, David C. H. Wallom
Oxford e-Research Centre
University of Oxford
Oxford, UK
{ramon.granell,david.wallom}@oerc.ox.ac.uk

Colin J. Axon
Institute of Energy Futures
Brunel University
London, UK
colin.axon@brunel.ac.uk

Introduction

Non-domestic buildings account for approximately 10% of UK greenhouse gas emissions [1] of which retail has the largest portfolio. The use of automated meter reading and building energy management systems is prevalent to monitor and evaluate energy use. However, the deployment of sub-metering within the building is less common. For facilities managers of companies with multiple stores, the detailed evaluation and comparison of different sites is becoming increasingly important requiring new and advanced analytical techniques not currently in common use. One such technique is automated clustering [2] to create groups that show similar behaviours, with the aim of understanding and finding unusual events. However, before applying these techniques, the general behaviour to be clustered should be represented in a meaningful and concise way. The temporal resolution of the profiles (from seconds to hours) has an impact on the clustering results [3].

Methodology

In an automated fashion, we obtain and exploit features from the daily electricity load profiles provided by a UK retail company. We demonstrate that a generalised set of features can represent the essential information of the consumption of UK retail establishments. The daily energy profile (the continuous line) in Fig. 1 is typical for a UK store. It presents a value for each one of the time intervals independent of there being a significant event *e.g.* night-time values can be considered constant. The single broad peak shown is directly related with their opening and working times.

Considering these characteristics, we propose to summarise the establishment's behaviour with eight features extracted from the profile: the start and end times of the peak, the start and end times of the stabilisation of the peak/off-peak periods *i.e.* t_0, t_1, t_2 and t_3, the average during the peak/off-peak phases *i.e.* $\mu(s_0)$ and $\mu(s_2)$, and the slope of the transition between the peak and off-peak periods *i.e.* $m(s_1)$ and $m(s_3)$. We can then compute an approximation to the real profile — the dashed line — and select the set of features with the smallest Euclidean distance between the approximated profile and the real profile.

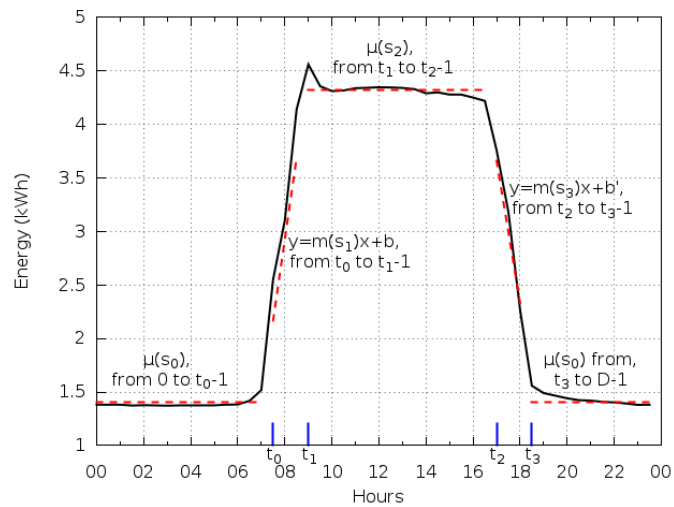

Fig. 1. Approximation to the real profile (continuous line) of a store based on the proposed features.

Results

We performed experiments on a dataset of more than 600 stores (from a single company) with electricity meter readings at thirty minute intervals acquired between April 2013 and October 2014. The eight proposed features were extracted from each store's Monday-to-Saturday profile. Then we applied a clustering algorithm based on the Gaussian Mixture Model [4] to group stores with similar behaviour. The resulting four clusters (using the slope values) are shown in Fig. 2. In this example, the number of stores in each one of the clusters is 84, 376, 64 and 118 for clusters 1, 2, 3 and 4 respectively.

Figure 3 shows the energy profiles representing the cluster centroids. There are clear differences in the magnitudes of the peaks for three of the four cluster centroids. Clusters 1 and 2 show similar consumption during the peak. However, the off-peak values of cluster 1 are significantly higher than in cluster 2. For these two clusters, there are also clear differences in the slopes.

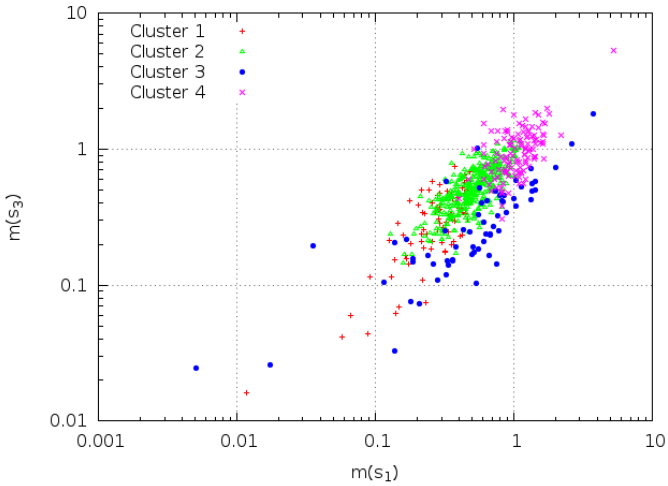

Fig. 2. Values of $m(s_1)$ against $m(s_3)$ depending on the store cluster.

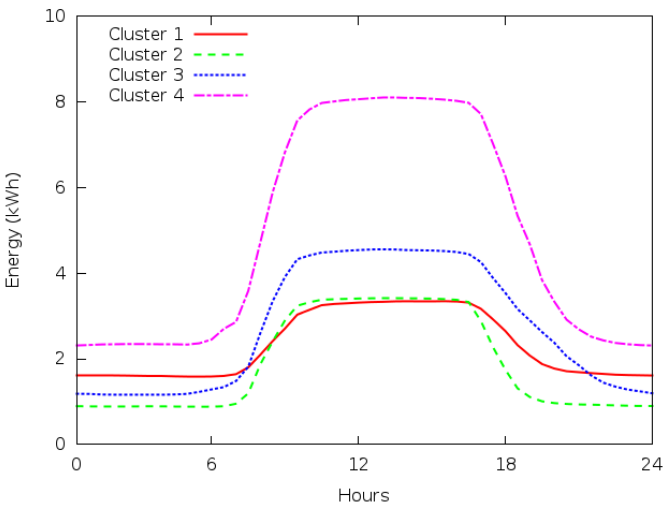

Fig. 3. The centroids of the obtained clusters.

Conclusions

We have presented a concise method to extract a set of features to represent the energy profiles of UK retail premises. Using these features we have clustered the stores and obtained distinct behavioural groups. As we represent each customer with a small number of features, combining features obtained from different profiles of the same customer *e.g.* gas/electricity, weekend/weekday, Summer/Winter, is relatively easy. Furthermore, the proposed features are independent of the temporal resolution of the profiles, allowing us to compare customers with different resolution profiles.

References

[1] WSBF and C. Connect, "Building efficiency: Reducing energy demand in the commercial sector," Westminster Sustainable Business Forum and Carbon Connect, Tech. Rep., 2013.

[2] G. Chicco, "Overview and performance assessment of the clustering methods for electrical load pattern grouping," *Energy*, vol. 42, no. 1, pp. 68–80, 2012.

[3] R. Granell, C. Axon, and D. Wallom, "Impacts of raw data temporal resolution using selected clustering methods on residential electricity load profiles," *Power Systems, IEEE Transactions on*, 2014, doi: 10.1109/TPWRS.2014.2377213.

[4] CountingLab, "Grouper 2.0 user guide, version 1," 2012.