The AAAI-17 Workshop on
Artificial Intelligence for Smart Grids and Smart Buildings
WS-17-05

Comparison of Clustering Techniques for Residential Energy Behavior using Smart Meter Data

Ling Jin, Doris Lee, Alex Sim, Sam Borgeson,
Kesheng Wu, C. Anna Spurlock, Annika Todd
Lawrence Berkeley National Laboratory 1 Cyclotron Rd, Berkeley, CA 94720
ljin@lbl.gov

Abstract

Current practice in whole time series clustering of residential meter data focuses on aggregated or subsampled load data at the customer level, which ignores day-to-day differences within customers. This information is critical to determine each customer's suitability to various demand side management strategies that support intelligent power grids and smart energy management. Clustering daily load shapes provides fine-grained information on customer attributes and sources of variation for subsequent models and customer segmentation. In this paper, we apply 11 clustering methods to daily residential meter data. We evaluate their parameter settings and suitability based on 6 generic performance metrics and post-checking of resulting clusters. Finally, we recommend suitable techniques and parameters based on the goal of discovering diverse daily load patterns among residential customers. To the authors' knowledge, this paper is the first robust comparative review of clustering techniques applied to daily residential load shape time series in the power systems' literature.

I. Background and Motivation

Cluster analysis is a commonly used unsupervised learning technique that can help identify different types of energy consumption behavior and has traditionally been applied to individual industrial and commercial customers or large aggregation of residential customers (Chicco et al. 2006 and Chicco 2012). The fast growing stream of interval meter data has motivated more recent research to apply such techniques to individual residential customers (e.g. Flath et al. 2012; Haben et al. 2016; Cao et al. 2013; McLoughlin et al. 2015; Rhodes et al. 2014).

Despite progress in clustering interval meter data for the residential sector, most aforementioned whole time series clustering was applied to customer-level data where each customer is associated with one selected or pre-aggregated load shape. In the context of demand response and efficiency programs, however, such an approach ignores the day-to-day differences within customers which relate to each customer's suitability to various demand side management strategies. For example, households with variable consumption schedules may be more likely to respond to time of use pricing incentives, whereas those with regular demand during the daytime are ideal for solar energy.

Clustering daily load profiles allows each customer to be represented by a number of representative load patterns and thus variability information can be derived. It can therefore serve as a valuable preprocessing step that provides fine-grained information on customer attributes and sources of variation for subsequent modeling and customer segmentation. Limited studies have explored daily load patterns within and across customers. Kwac et al. (2014) clustered individual household-day load profiles and found that although two homes might have the same average profiles, the "information entropy" or diversity of load patterns from one day to the next could vary significantly. McLoughlin et al (2015) differentiated customer profile classes by first clustering their day-to-day usage patterns.

Clustering results are known to be highly sensitive to the nature of the data and choices of algorithm, parameter settings, and data cleaning strategies (Luxburg U 2012). The validity of clustering algorithms needs to be assessed for different types of data. Existing work in the area of clustering load profiles includes comparative studies of clustering techniques and the application of novel clustering methods, including Adaptive KMeans (Kwac et al. 2014), Spectral Clustering (Vercamer et al. 2015), Self-Organizing Maps (Verdu et al. 2006), and subspace clustering (Piao et al. 2014). Kim et al. (2011), McLoughlin et al. (2015), Hsu (2015), and Chicco et al. (2006; 2012) conducted comparative studies to assess the performance of different clustering algorithms. However, most of these studies have based their assessment on small, non-residential dataset (<1000 load curves). Residential customers are characterized by highly volatile behavior, which challenges the application of clustering methods to individual load curves (Chicco 2012). Because of the size of future Advanced Metering Infrastructure (AMI) datasets scalability is also a concern in the choice of clustering algorithm. In addition, Piao et al. (2014) observed that distance-based clustering is often ineffective in the high dimensions of time-series usage profiles, as most data points are equidistant and hence indistinguishable from one other. To address these concerns, our contributions are to: 1) assess clustering algorithms on a large number (10^5) of individual residential daily load curves; 2) compare 11 different algorithms using 6 domain-specific and generic comparison metrics, 3) assess whether density-based and probabilistic methods perform better for this high-dimensional AMI dataset than distance-based methods.

The paper is structured as follows. Section II introduces the interval meter data we use, clustering algorithms we test, and performance metrics we consider. Section III explores the parameter space and recommends the optimal parameter choices. Section IV conducts a cross comparison among all clustering algorithms and discusses their strengths and weaknesses. Section V concludes and recommends directions for future work.

II. Data and Methods

We focus on clustering daily consumption patterns based on the timing and relative magnitude of individual households' discretionary electricity usage instead of absolute consumption. The clustering goal is not to forecast hourly consumption load but rather to identify a set of typical consumption behaviors that differ in their timing throughout the day. The hourly metering data are obtained from a summer peaking utility in California, and consist of over 30 million daily load profiles from approximately 100,000 households, with a monitoring period spanning from June 1st 2011 to May 31st 2012. To achieve robust results, we first clean the data and decide the format of the object to be clustered. Then we apply a suite of clustering techniques and explore the suitable choice of parameters for each technique based on cluster performance metrics for compactness and/or distinctness. Finally, we compare the clustering techniques and discuss strengths and weaknesses.

Data Cleaning and Normalization

Daily usage data with missing hours (0.3% of the data) or with very small power demand (<0.2 kW on average; 6% quantile) are ignored in populating our clusters. These days are either affected by blackouts, meter malfunction, or no occupancy of the household. After data cleaning, 32,611,421 daily load shapes (94% of the dataset) remain for the full year period.

Our interest being primarily the temporal aspect of the daily profile rather than absolute usage, the load profiles are preprocessed with normalization. Most existing studies cluster based on normalizing the daily usage data by a reference power following standard methods reviewed in Milligan and Cooper (1998), such as normalization by daily maximum (Chicco et al. 2006; Chicco 2012), min-max normalization (Piao 2014; Han et al. 2012; Cao et al. 2013), and normalization by daily total (Kwac et al. 2014).

We propose and apply a new normalization scheme following Jin et al. (2016) that focuses on characterizing timing and relative magnitude of discretionary consumption. Specifically, daily minimum demand is subtracted from hourly usage (de-minning) and each hour is then divided by the de-minned total. Daily minimum power demand serves as a proxy for "baseload;" after normalization, a load shape represents its hourly contribution to daily total discretionary usage and shape clusters can be interpreted in terms of timing of higher and lower discretionary demand.

Clustering Algorithms

A random subsample of 325 households (104673 daily load profiles) with data over the full year are used to evaluate a suite of clustering techniques in four categories.

1. Centroid-based Methods are a class of algorithms that iteratively assign and update each observation to its closest centroid, which can be defined as the mean or median. This can be formulated as an expectation-maximization problem and the iteration terminates when the results converge (centroids remain unchanged).

The KMeans (Anderberg 1973) algorithm seeks a good cluster by minimizing the within-cluster sum of squared residuals. The KMedoids algorithm is similar to KMeans, except KMedoids takes the cluster centroid as the medoid of the data points. Therefore, KMedoids guarantees that the cluster centroid itself is among the cluster members. The motivation for using medians as centroids over means is that medians are less skewed by outliers, so KMedoids has the potential to be more robust.

Adaptive Kmeans (AKmeans), developed by Kwac et al. (2014), combines partitioning and hierarchical algorithms. The algorithm first partitions the dataset into a large number of clusters such that the relative squared error (RSE) of any load shape assigned to a cluster is not greater than an error threshold θ. Then the clusters are hierarchically merged by sequentially combining the most similar clusters until their total count reaches a target number. The user can define the violation rate of clusters that exceed the error threshold θ to ensure the quality of the clusters.

2. Hierarchical Clustering is a family of algorithms that takes an agglomerative or divisive approach to build a hierarchy of clusters. It has been widely applied in load shape clustering as reviewed in Chicco et al. (2006; 2012). The algorithm uses a linkage criterion to determine the distance between different sets, as well as a distance metric for computing the similarity between pairs of data points. In Chicco (2012), findings suggest that a non-Euclidean distance metric and certain linkage methods might perform better than the L2 norm. We explore a number of linkage criteria in this study (Ward's, average, and complete linkage) along with two distance metrics (Chebyshev and Euclidean).

3. Density-based Clustering was proposed by Ester et al. (1996) when they observed that distance-based clustering was often ineffective in large size datasets with high dimensions, such as time-series usage data where each dimension was assigned with equal weights. DBSCAN (Ester et al. 1996) is a density-based method that visits every data point and finds other data points that lie within some chosen epsilon distance. After getting all points in the epsilon neighborhood, if there is more than a user defined minimum number of data points in the neighborhood then this forms a core, if not then this data point is considered noise. The cores are then recursively merged into larger clusters. This density-based approach often enables DBSCAN to find non-linear shaped clusters better than other algorithms, which can be good for load profile applications.

4. Model-based Clustering is based on fitting a probability distribution over the clusters. Since they are a generative process, the clustering is independent of distance metrics which may be useful for our high dimensional dataset. This study applies Gaussian Mixture Model (GMM) to load shape clustering. GMM fits a number of Gaussians to the data. The algorithm is motivated by the observation that the products of many Gaussians are still Gaussians. Therefore, by increasing the number of components of Gaussians and finding suitable parameters (means and covariance), we can build a better representation of the data.

Cluster Validity Metrics and Criteria

The metrics used for comparing different model parameters are based on cluster geometry: a good cluster should be very "compact" (its elements are very close to one another) and "distinct" (the center of a cluster should be far away from center of another cluster). The metrics are chosen such that ground truth labels are not required to compare the clustered values. While there are many variations of these indices in the literature, they are modified here for interpretability so that the lower the index the better the clustering. Euclidean distance is the metric we used for comparison. Two common measures of cluster geometry are infra-set distance of set *S*, which measures the pairwise distances between the points, and *scatter* with respect to the cluster centers. Many of these clustering indices were also used in (Dent 2012) except for the violation rate of a RSE threshold (VRSE). The VRSE (0.3 recommended in Jin et al. 2016) measures how close the data lie to the centroids. This metric was used in Kwac et al. (2014) to ensure the typical load shapes derived from cluster centroids are representative of their cluster members within a controllable error threshold. The clustering indices used in this study are detailed in Table 1.

Lastly, besides validity metrics, clustering techniques need to be compared by their suitability to segmentation goals using post-cluster checking. For example, Chicco (2012) found that clustering validity metrics tended to favor methods that resulted in highly uneven clusters and thus can only be used to isolate outliers. Our segmentation goal is to identify a diverse set of typical daily shapes that can be described by the cluster centroids representing different patterns in consumption schedules within and across customers. Criteria based on this goal will be proposed and discussed after cross comparison of the validity metrics.

III. Application of Clustering Algorithm

The performance of different types of clustering algorithms depends on the choice of their respective parameters and variants. An optimal choice can be revealed by examining the evolution of the performance of indices over the parameter space. Previous reviews of load shape clustering focused on deriving customer classes each represented by certain load shapes and therefore generally limited the number of clusters to 10-15 (Chicco et al. 2006; 2012; Cao et al. 2013). This limited number of clusters is often not mathematically adequate, but is practically necessary for tariff purposes. However, in this study, instead of identifying different types of customers, we focus on clustering daily load profiles to reveal diverse consumption behaviors within and cross customers and therefore relax the search of the optimal number of clusters to be well beyond 15.

Centroid-based Methods

Centroid-based clustering can yield non-deterministic results due to random initialization, therefore we have cited the results from running the algorithm over 3 random initializations. The key parameter in Kmeans and Kmedoids algorithms is the target cluster number. The optimal number of clusters is decided by examining the "elbow" location of clustering indices as a function of cluster numbers ranging from 10 to 500. We find the optimal number of clusters occur around 90 for Kmeans and 50 for Kmedoids.

AKmeans involves more iterations and therefore is computationally more expensive. The resulting number of clusters from the partitioning procedure depends on the error threshold (θ). We follow Jin et al. (2016), which found $\theta = 0.3$ was a good choice based the criteria that the number of clusters be not large (~5000) and the marginal

Table 1 List of clustering index used for comparison

	Equation	Description	Measure
CDI	$CDI = \frac{1}{\hat{d}(C)}\sqrt{\frac{1}{K}\sum \hat{d}^2(R_k)}$	Cluster Dispersion Indicator	compactness and distinctness
MIA	$MIA = \sqrt{\frac{1}{K}\sum_{K=1}^{K} d^2(r^{(k)}, C^k)}$	Mean Index adequacy	compactness
Silhouette	$SIL = \frac{max(a,b)}{b-a}$ where a = average intra-cluster distance, b = average shortest distance to another cluster	Inverse Silhouette index $\epsilon[-1,1]$. If SIL¡0, cluster not very compact. Note that this is the inverse of typical definitions of SIL in literature.	compactness and distinctness
Average SMI	$\alpha_{ij} = \frac{1}{1-\frac{1}{ln[d(C_i,C_j)]}}$ $< SMI >= \frac{1}{N}\sum_i \sum_j \alpha_{ij}$	Similarity Matrix Indicator generates a KxK matrix, where K is number of cluster. The farther away the non-diagonal elements are from 1 the better. quantify the measure, averaging over the whole matrix gives us a sense of how non-diagonal elements behave.	distinctness
DBI	$DBI = \frac{1}{K}\sum max \frac{scatter(C_i)+scatter(C_j)}{d(C_i,C_j)}$ where $i \neq j$	Davies-Boulden indicator	compactness and distinctness
VRSE	$E(s, i*(s)) = \sum(s(t) - C_i^*(t))^2 \leq \theta \sum C_{i^*_{(s)}}(t)^2$	Violation rate of RSE threshold. Percentage of data that lie beyond a threshold distance away from the . centroid. With $\theta = 0.3$	compactness

gain in error improvement to explanatory power is small. By limiting the total violation rate to 5%, the hierarchical clustering merged the akmeans clusters to 2000 clusters. Following Kwac e al. (2014) we further truncate the cluster centers with lowest membership counts so that the violation rate is 30%, 40%, 50%, 60% and 70%, which reduces the 2000 clusters to 96, 52, 24, 13, and 5 respectively.

Hierarchical Clustering

Hierarchical clustering first computes a NxN similarity matrix (called a linkage) that describes how far away data points are from each other. We employ a bottom-up, agglomerative approach for merging the closest clusters and continue to do this until there is only one large cluster left.

The similarity matrix depends on choices of distance metrics and linkage methods. Chicco (2012) reviewed hierarchical methods applied to non-residential data and found non-Euclidean distance metrics and certain linkage methods might perform better than the L2 norm. To examine which distance and linkage method is best for our dataset, we perform an exploratory analysis based on the Cophenetic correlation coefficient (CCC), which measures the linear correlation between the pairwise distances and the dendrogram distances between two data points.

$$c = \frac{\sum_{i<j}(Y_{ij} - \langle Y \rangle)(Z_{ij} - \langle Z \rangle)}{\sqrt{\sum_{i<j}(Y_{ij} - \langle Y \rangle)^2 \sum_{i<j}(Z_{ij} - \langle Z \rangle)^2}} \quad (1)$$

Where Y_{ij} is the pairwise distance matrix and Z_{ij} is the cophenetic distance between points *i,j*. The closer the CCC value is to 1 the better the dendrogram preserves the distances between the clustered and raw data points. With a range of different distance metrics and linkage methods, we find that complete and average linkages along with Chebyshev and variants of Euclidean distance yield the highest Cophenetic coefficient (Figure 1). Consequently, the subsequent analysis is based on Chebyshev and Euclidean distance metrics.

Linkage criterion is the strategy that the algorithm uses for merging the clusters. Ward's linkage agglomeratively adds in points to minimize the within-cluster sum of squared residuals. Average linkage criterion uses the average of all pairwise distances between all data points from two different clusters as its objective function to merge the nearest clusters. Complete linkage merges the clusters based on the farthest distance between two clusters.

To get meaningful clusters out of the similarity matrix, we specify the number of clusters and go back to the merging history, to see which split levels resulted in the K clusters. This could be thought of as making a cut through the dendrogram. The optimal number of clusters are generally found around 20 to 30 among all the hierarchical variants when a relatively clear “elbow” location can be identified in the index-to-cluster-number curve.

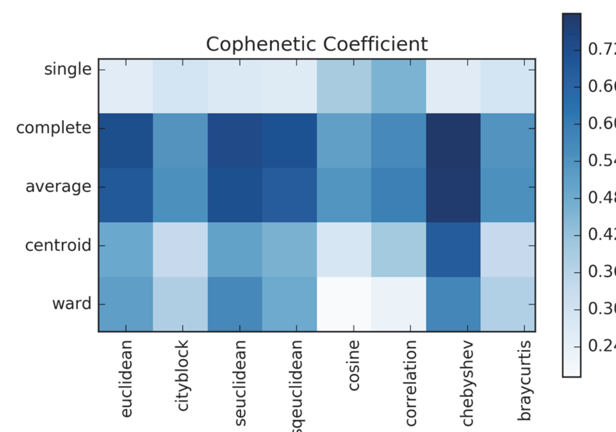

Figure 1 Cophenetic coefficient computed by sampling on 20% of the data.

Density-based Clustering

DBSCAN is sensitive to parameters epsilon and minPt, which determine how many clusters the data are segmented into. To estimate epsilon, we use a Ball tree to find the distance from each point to its nearest neighbors. To estimate minPts, we query the number of points within the epsilon neighborhood of every other point. We are only able to obtain an order of magnitude estimate that the minPts should be small (< 100). We use these coarse range estimates to conduct several iterations of grid search that yielded a suitable number of clusters. The final ranges used for comparison in Figure 2 is epsilon 2-6 and MinPt 0.15-0.21. Figure 2 indicates the optimal number of clusters is determined by the elbow around 50. However, the number of clusters we are able to tune depends on a two-dimensional parameter space and therefore is fairly limited compared to other clustering methods so the exact optimal number is not certain. In the end, the majority of the data are classified as noise (>90%), and the number of daily load shapes that get clustered are very small (<10%).

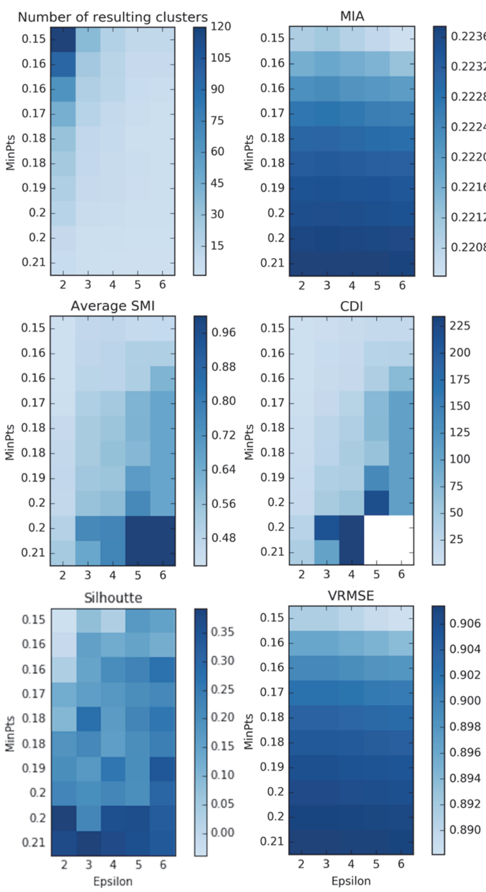

Figure 2 Cluster number and clustering indices as a result of grid search with DBSCAN in epsilon and minPts.

Model-based Clustering

The GMM employs an EM algorithm to find a suitable number of clusters based on the Bayesian Information Criterion (BIC). After model-fitting process, we extract the cluster centroids as the mean of the Gaussian components.

The AIC (Akaike Information Criterion) and BIC for different numbers of components is plotted in Figure 3, revealing the optimal number is around 15 to 20.

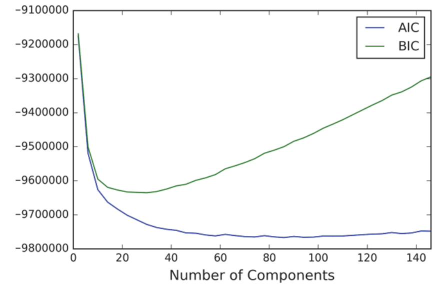

Figure 3 Information criterion for different numbers of components in the GMM Model.

IV. Method Cross Comparison

The comparison presented in this section is based on clustering results from different methods with numbers of clusters varying from 10 to 150. Hierarchical methods are labeled after their respective linkage criterion (Ward, Average, or Complete) and distance metric (Euclidean or Chebyshev). Cluster validity metrics are computed and presented as a function of number of clusters in Figure 4. Among these metrics, MIA and VRSE measure the "compactness" of the resulting clusters; SMI measures the "distinctness"; and CDI, Silhouette, and DBI measure both. The lower the metric value the better the clustering.

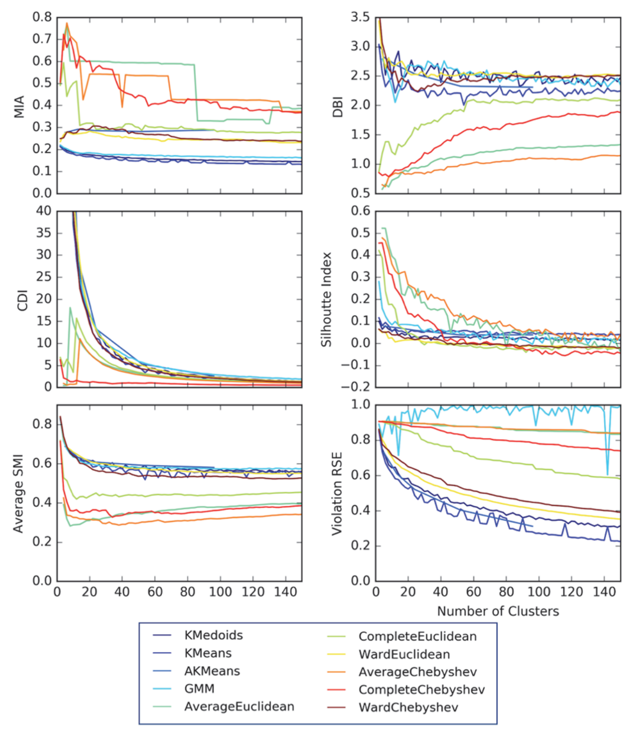

Figure 4 Validity metrics varying with the number of clusters across algorithms.

Figure 4 indicates that the performance ranking of various methods is not consistent across the validity metrics and there is a trade-off between their ability to capture "compactness" vs "distinctness." In general, centroid methods and hierarchical methods with Ward linkage criteria tend to produce compact clusters as indicated by lower MIA and VRSE indices, while hierarchical methods with average or complete linkage criteria are better at producing clusters that are distinct from each other. The GMM method produces relatively compact clusters as indicated by a low MIA index, however, a high violation rate (>90%) of RSE suggests that its resulting centroids (i.e. cluster means) are not representative of their respective member load shapes. Validity metrics (CDI, DBI and Silhoutte index) that measure both compactness and distinctness tend to favor methods that produce more distinct clusters. Also noted is that the performance of hierarchical clustering methods does not always improve with increasing numbers of clusters except for Ward linkage criteria.

Our findings are clearly different from Chicco (2012) where performances of various methods were found to be clearly differentiated by the cluster validity metrics. In Chicco (2012), centroid based methods were found to perform the worst while some variants of the hierarchical methods performed the best. This difference may be due to the data Chicco use; which were obtained from non-residential sectors that are less noisy and are much smaller in size (400 load shapes instead of 10^5 used here). Additionally, the normalization method used in Chicco was min-max based, while we aim instead at deriving hourly contributions to discretionary usage. The min-max normalization test on our dataset, however, reveals significant amplification of noise in the load patterns and consistently worse performance across the validity metrics.

As an unsupervised learning problem, there may be different cluster methods for different purposes. When clustering performance exhibits a clear tradeoff between "compactness" and "distinctness," the choice of preferred clustering method needs to be examined based on suitability to segmentation goals. As described earlier, clustering daily load shapes here is intended as a preprocessing and data reduction step as a basis for subsequent customer-level feature extraction and segmentation. More specifically, the goal is to identify a diverse set of typical daily shapes that can be described by the cluster centroids representing different patterns in day-to-day and customer-to-customer consumption schedules. To satisfy this goal, two aspects of the clustering results are preferred:

- High statistical affinity of cluster centroids to their member load shapes, and
- Avoiding highly uneven cluster sizes.

The first aspect ensures that the resulting typical daily shapes are representative of actual consumption schedules and thus can be used to describe the time of use patterns of their respective cluster members. The second aspect ensures a diverse set of daily patterns are identified each with adequate data coverage.

Post-cluster checking indicates that the above two aspects are highly correlated in the clustering results. Figure 5 shows an example of the cluster size distribution resulting from various methods for a total cluster number (K) of 50. The clustering methods in Figure 5 (from top to bottom) are sorted by the Chi-square statistic (from low to high), against a uniform distribution. A lower Chi-square statistic indicates more evenly distributed cluster sizes. The centroid based methods and Ward hierarchical method tend to produce compact clusters (Figure 4) with more evenly distributed sizes (Figure 5). Consequently, these methods are clearly preferred. Other methods, despite better performance indicated by distinctness and some validity measures, fail to produce compact clusters and the centroids are not representative of their load shape members (measured by VRSE). In fact, the "distinctness" is achieved at the expense of a result including a single large and noisy cluster that often account for >80% of the data. "Compactness" is therefore a more valuable quality of preferred clustering technique applied to our dataset.

For DBSCAN, 99% of the data are classified as noise and excluded from clustering. Appropriate application of density based methods may require further preprocessing (such as smoothing) of the highly variable residential data.

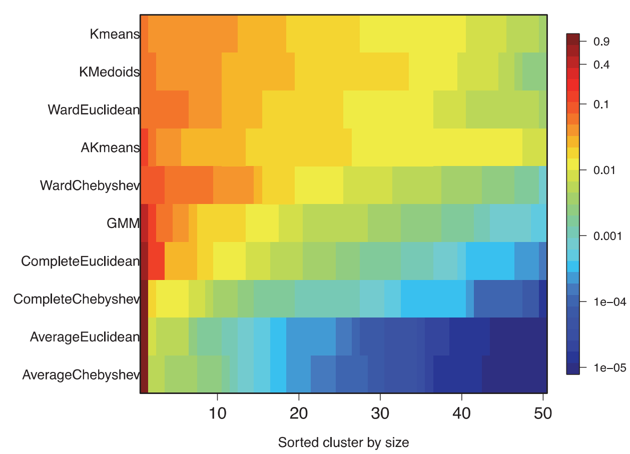

Figure 5 Fraction of load shapes assigned to each cluster for K=50. Clustering methods are sorted by chi-square test statistic against uniform distribution.

V. Conclusion and Outlook

We have evaluated the performance of 11 clustering methods under 4 families of algorithms: centroid based, hierarchical, density based, and model based methods. Whole time series clustering is examined with more than 10^5 daily residential load shapes that are processed to focus on individual households' discretionary electricity usage patterns. The parameter settings are evaluated for individual algorithms to determine best choices unique to the residential daily load shapes using six performance metrics.

The performance ranking of various methods is not consistent across the validity metrics and there is a trade-off pattern between their ability to capture "compactness" vs "distinctness" in the resulting clusters of our dataset, which is then resolved by post-cluster checking guided by our segmentation goal. The goal of identifying a diverse set of typical consumption schedules requires resulting clusters are compact and relatively evenly distributed. Post-cluster examination reveals that algorithms with heuristics minimizing the within cluster scatter, i.e. centroid based methods and Ward linkage hierarchical methods, perform better with respect to the segmentation goal.

Density based methods classify the majority of our residential dataset as noise with very little data being clustered. Future application of this type of method should include smoothing of the load data in the preprocessing step for noise reduction and/or dimension reduction such as clustering on a reduced number of features derived from the raw load data.

Model based methods generally perform in the middle among all the methods examined here in terms of producing compact and evenly distributed clusters. However, the centroids as the mean of the Gaussian components fail to represent the temporal patterns in their respective member load shapes.

Our results highlight the limitations of using a single clustering validity metric to guide the selection of clustering methods when there is a clear tradeoff between compactness and distinctness. When load shape clustering is intended as a preprocessing step for subsequent household-level segmentation, the requirement for low numbers of resulting clusters, previously practical for tariff purposes, should be relaxed in order to capture the diverse time of use behavior in residential daily consumptions. In addition to the size distribution and visual examination conducted in the post-cluster checking in this study, broader validation of the robustness of the identified clusters needs to be evaluated with additional data sources, such as household demographic and socio-economic information, rate structure, weather, etc. In particular, the sources of variability in the identified daily patterns need to be established and better understood to support more effective demand side management and behavior based programs.

Acknowledgments

This work was supported by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

References

Ankerst, M. et al. 1999. OPTICS: Ordering Points To Identify the Clustering Structure. SIGMOD '99 Proc. of the 1999 ACM SIGMOD Int. Conf. on Manag. of data, 28, no. 2, pp. 49–60.

Cao, H. et al. 2013 Are domestic load profiles stable over time? Proc. 39th Annu. Conf. IECON, pp. 75-86.

Chicco, G et al. 2006. Comparison among clustering techniques for electricity customer classification. 21(2), pp. 1–7, 2006.

Chicco, G. 2012 Overview and performance assessment of the clustering methods for electrical load pattern grouping. Energy, 42(1), pp. 68–80, 2012

Dent, I et al. 2012. An Approach for Assessing Clustering of Households by Electricity Usage. The 12th Annual Workshop on Computational Intelligence, pp. 5–10.

Ester, M et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. pp. 226–231, 1996.

Haben, S et al. 2016. Analysis and clustering of residential customers energy behavioral demand using smart meter data, IEEE Trans. Smart Grid, 7(1), pp. 136-144.

Han, J et al. 2012. Data Mining: Concepts and Techniques.vol. 3.

Hong, W. C. 2011. Electric load forecasting by seasonal recurrent SVR with chaotic artificial bee colony algorithm. Energy 36(9): 5568-5578.

Jin, L., et al. 2016. Load shape clustering using residential data. To be submitted.

Kwac et al. 2014. Household energy consumption segmentation using hourly data. IEEE Trans. on Smart Grid, 5(1), pp. 420–430.

McLoughlin, F et al.2015. A clustering approach to domestic electricity load profile characterisation using smart metering data. Applied Energy, vol. 141, no. March, pp. 190–199.

Moslehi, K. and R. Kumar. 2010. A Reliability Perspective of the Smart Grid. IEEE Transactions on Smart Grid 1(1): 57-64.

Piao, M et al. 2014. Subspace Projection Method Based Clustering Analysis in Load Profiling. IEEE 29:6, pp. 2628–2635.

Rhodes, J. D. et al. 2014. Clustering analysis of residential electricity demand profiles. Applied Energy 135: 461-471.

Sculley, D. 2010. Web-scale k-means clustering. Proc. of the 19th international conference on World wide web. 10, p. 1177, 2010

Vercamer, D, B. Steurtewagen, D. Van den Poel, and F. Vermeulen. 2015. Predicting Consumer Load Profiles Using Commercial and Open Data. IEEE Trans. on Power Systems, pp. 1–9, 2015.

Verdu, S.V., et al. 2006. Classification, filtering, and identification of electrical customer load patterns through the use of self-organizing maps. IEEE Trans. on Power Systems, 21(4), pp. 1672–1682.

VonLuxburg U, Williamson B, Guyon I. 2012. Clustering: science or art? J Mach Learn Res 27:65–80.