A Robust Unsupervised Framework for High-Resolution Building Energy Consumption Profiling

Sicheng Zhan[1], Adrian Chong[1]
[1]Department of Building, School of Design and Environment,
National University of Singapore, Singapore

Abstract

Unsupervised learning methods have been widely used for building energy consumption profiling, but the currently used methods usually gave undesirable results and could hardly tolerate the highly diversified building cases. A scalable automatic framework is accordingly proposed in this study to achieve accurate load profiling. The framework consists of four major steps: pre-processing, preliminary K-means clustering, DBSCAN within clusters, and post-processing. With a dataset including 50 different buildings in Singapore, the framework was demonstrated to outperform all the baseline methods in most cases (37 out of 50). The profiling results provides more comprehensive insights on the buildings energy behavior, facilitates applications such as building load prediction and improves the prediction accuracy.

Introduction

With the escalating deployment of smart meters and building management systems during the past decades, a huge amount of building operational data becomes available. Building level temporal energy consumption is the most common data. It was realized that traditional statistical or physical principle-based methods were not capable of fully exploiting the information embedded in these high-dimensional and highly diversified data (Fan et al., 2018). Thus, various advanced data mining techniques have been introduced to analyze and apply the data for building energy assessment and building operation.

Load profiling and its application

One important application of building temporal energy consumption data is energy consumption profiling (also called load profiling), to identify the representative diurnal energy usage patterns of a building. The profiling result provides designers, engineers, and facility managers with better knowledge on building operation. It is also useful for many further applications, including but not limited to building energy simulation, occupancy and load prediction, demand side management, and abnormal operation detection.

One application of load profiling is to classify the customers according to the identified profiles. The classification was done both at occupant level for domestic buildings (Tsekouras et al., 2007; McLoughlin et al., 2015) and at building level for non-domestic buildings (Li et al., 2018). With the identified representative profiles, some demand side management approaches were applied to achieve load shape adjustment objectives such as "peak clipping" and "valley filling" (Wei et al., 2018; Panapakidis et al., 2014). Fault detection and diagnosis was done both in the process of profiling (Jalori and T Agami Reddy PhD, 2015a) and according to the profiling result (Habib and Zucker, 2015). Also, the profiling result was used to assist in building energy consumption prediction. By fitting prediction models separately for different typical profiles, both the accuracy and the efficiency were improved by Tang et al. (2014) and Shahzadeh et al. (2015). And for physical-based building energy models, where occupant behavior was claimed to account for up to 30% of the uncertainty (Eguaras-Martínez et al., 2014), the profiling result can be used to infer the schedules as inputs of the model.

Existing methods for load profiling

Buildings' characteristics tend to be very different from each other and sufficient background information is hardly available. Therefore, though other methods including regression and neural network have also been used for load profiling, unsupervised learning methods were recognized to be more promising (Miller et al., 2018). Studies have used unsupervised learning methods for load profiling, including K-means clustering (Tsekouras et al., 2007), Hierarchical clustering (Fan et al., 2015), Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Jalori and T Agami Reddy PhD, 2015b), K-shape clustering (Yang et al., 2017), Self-Organizing Map (Panapakidis et al., 2014), and Symbolic Aggregate approXimation (Miller et al., 2015). Next, we will discuss K-means and DBSCAN in detail as they represent the two most popular categories of clustering method: distance-based and density-based partitional clustering. We will also cover the comparison and analysis of other typical methods later in the discussion section.

Among the diverse clustering methods, K-means was most widely investigated. Hsu (2015) showed that it gives the most stable clustering result if the cluster number K is correctly chosen. Green et al. (2014) showed its effectiveness in facilitating the subsequent prediction. By carefully picking the initial centroids and combined with Self-Organizing Map, Panapakidis et al. (2014) further improved the clustering accuracy. However, K-means assumes that the clusters have spherical variance and are of similar size. Consequently, K-means tends to wrongly cluster non-spherical or unevenly sized clusters. Moreover, it requires a presumption of the cluster number, which is usually questionable and leads to the ignorance of special profiles on minor certain days. Shown in figure 1.a is an example where K-means failed to correctly recognize the elongated clusters from a building's diurnal load data. The samples that were wrongly clustered are highlighted in the black dashed rectangles. It can be told from the lower plot that the correct clusters should be as in the dashed ellipses of the corresponding colors. This kind of problems happens more often in the case of load profiling because the samples are usually at high dimension (24 or higher), which is also known as "Curse of Dimensionality" (Verleysen and François, 2005) and will be further discussed later.

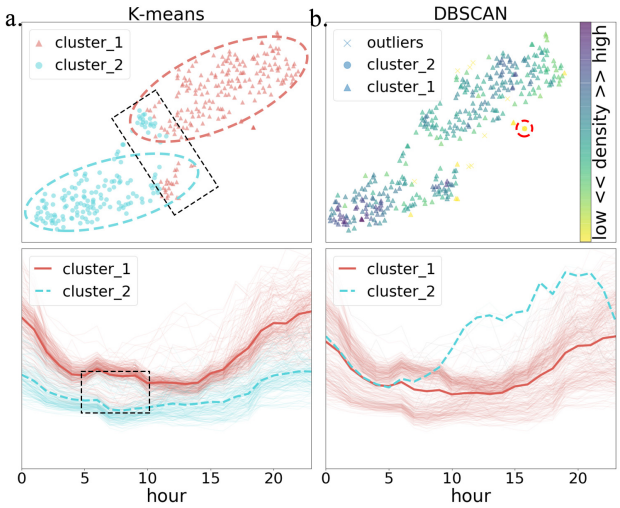

Figure 1: An example of (a) K-means and (b) DBSCAN giving bad clustering results (Upper: 2-D visualization; lower: time series load profile.)

By contrast, Ester et al. (1996) introduced DBSCAN as an alternative method that recognizes elongated clusters, identifies outliers and doesn't require predefined cluster number. For example, Jalori and T Agami Reddy PhD (2015a) applied DBSCAN to identify the essential diurnal schedules from the building energy interval data. However, it is still arbitrary and troublesome to tune the parameters, Epsilon and minimum points. Additionally, with a pair of parameters defining only one density threshold, DBSCAN is naturally unable to well identify clusters and outliers when the density varies over the samples. Figure 1.b displays the DBSCAN clustering result of the same building. The clusters are shown as marker types in the scatter and color in the plot, while color in the scatter plot stands for the density level around the samples. As highlighted in the red dashed circle, these few samples were relatively far from the others and of lower density, so the model identified them as a cluster and failed to separate the two major clusters.

Objectives

The objective of this study is to tackle the following three major problems of the currently popular methods for building electricity load profiling:

1. A single type of model was not suitable for varying characteristics of different buildings' energy data;
2. No existing method was able to well capture the detailed building energy usage behavior;
3. The models usually required intuitive tuning before being applied to a new building, contradicting the concept of unsupervised learning.

Therefore, we designed and implemented an automated framework to achieve high-resolution electricity load profiling for different buildings. In the following sections, we will illustrate the structure of the framework in detail. Then we will benchmark the proposed method against K-means, DBSCAN, Dynamic Time Warping (DTW) (Sakoe and Chiba, 1978) and K-shape (Paparrizos and Gravano, 2015). Also, we will use the clustering result to infer the operating schedule of the buildings to demonstrate its effectiveness for further application. Finally, we will discuss the results and directions to further improve the load profiling performance.

Methods

The proposed framework

As in figure 2, the framework consisted of 4 main steps: pre-processing, K-means clustering, DBSCAN clustering, and post-processing. With K-means and DBSCAN complementing each other, the hierarchical framework overcame their respective disadvantages. The Greedy algorithm was applied over the whole framework to tune the parameters automatically and efficiently. The Calinski-Harabasz (CH) index stood out as the objective from the various clustering validation indices (CVI) because of its higher robustness (Maulik and Bandyopadhyay, 2002).

As the first step, the raw annual hourly energy consumption data was filtered, normalized and reshaped into diurnal electricity load data. Considering that load shift on a single hour can make a profile different, days with any missing value were filtered off. Maximum normalization was applied to maintain information about the base load, which is important for understanding how a building is operated. After this, the 24-dimension load data of each day during the year was used for further clustering.

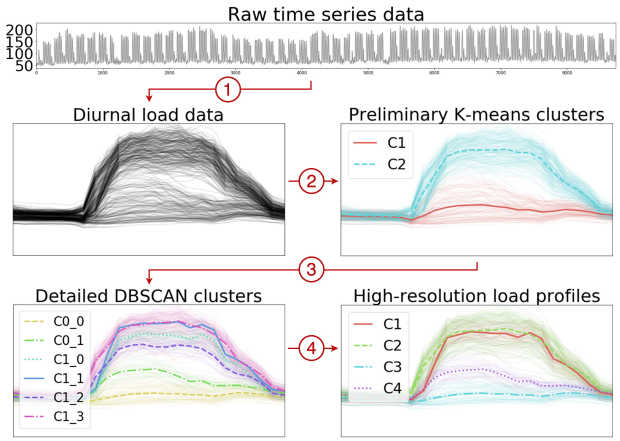

Figure 2: Information flow of the proposed framework (1.Pre-processing; 2.K-means; 3.DBSCAN; 4.Post-processing).

Afterwards, K-means was applied to initialize the clustering because it is reliable to give rough clustering result (Hsu, 2015). Based on the nature of building energy usage, the cluster number K was optimized between 2 and 10 against the CH index. However, K-means has two major issues for the task of load profiling: a) failure in recognizing detailed building energy behavior, and b) wrong clustering of samples at the joint area between two elongated clusters. These two issues were addressed in the following steps.

As the third step, DBSCAN was used to cluster within the preliminary clusters obtained by K-means. The number of minimum points was selected as 2 so that any pattern happening on more than two different days is recognized as a cluster. And the other parameter Eps was tuned between 0.02 and 0.6 against the CH index. The outliers were treated as individual clusters by themselves when calculating the index. In this way, different parameters were applied on different preliminary clusters, so that the inherent weakness of DBSCAN in clustering data with different density was eliminated. Additionally, its ability to identify fine clusters and outliers compensated for the first issue of K-means. This was the most critical step to obtain high-resolution building energy load profiles.

By post-processing, the second issue of K-means was solved and the final clusters were identified. The wrongly clustered samples were separated from the major part of its preliminary cluster when applying DBSCAN, and at this step combined into their true clusters. Pearson Correlation Coefficient (PCC) matrix of centroids (mean of all the samples in each cluster) was first calculated. Then starting from the highest PCC, clusters with PCC higher than a threshold were merged. Again, the threshold of PCC was optimized between 0.8 and 1 by comparing the CH scores of the final clusters.

Clustering result benchmarking

To demonstrate the effectiveness of the proposed framework, we tested the algorithm on a dataset containing a year of hourly building electricity consumption data from 50 campus buildings in Singapore. Building types included office, educational, residential and commercial. After preprocessing, the buildings had 357 days of data on average, 82% of the buildings had more than 350 diurnal electrical load data and 90% had more than 340. The clustering result was benchmarked against four typical clustering methods: K-means with Euclidean distance, DBSCAN, K-means with DTW, and K-shape. K-means and DBSCAN were selected as they are the most popular methods during the past decades, while the other two shape-based methods were claimed to be more suitable for time series clustering. Parameters of all these methods were also optimized against the CH index. We benchmarked the methods in three approaches: visual comparison, using CVI and schedule inference. The entire process was implemented in Python, using 'Pandas', 'Scikit-learn', and 'Tslearn' packages.

First, all the clustering results were visualized as in figure 1 to observe whether the clusters are well separated. T-Distributed Stochastic Neighbor Embedding was applied to map the 24-dimensional dataset to the 2-dimension scatter plot so that the spatial relationship between samples could be easily noticed. The centroids of each cluster were highlighted in the plots as bold lines to distinguish the different energy usage patterns such as early start or late finish. We will use overall evaluation and typical cases to illustrate how the proposed method outperforms the others.

Since the real buildings' operating schedule, which is the ground truth of clustering, is usually unavailable, external CVIs are not applicable. Therefore, we applied internal CVIs to evaluate the performance of the methods. However, most existing CVIs have their limitations and none has been proved to be robust, especially for high-dimensional datasets with diversified variance and outliers (Liu et al., 2010). Thus, both CH index and another mostly used CVI Silhouette Coefficient were calculated to compare the methods and to analyze the effect of CVIs. The CH and Silhouette scores are defined as equation 1 and 2, the meaning of which were clearly elaborated by Desgraupes (2013).

$$CH = \frac{N-K}{K-1} \frac{\sum_{k=1}^{K} n_k \left\| C^{\{k\}} - C \right\|^2}{\sum_{k=0}^{K} \sum_{i \in I_k} \left\| M_i^{\{k\}} - C^{\{k\}} \right\|^2}. \quad (1)$$

$$Silhouette = \frac{1}{K} \sum_{k=1}^{K} \left(\frac{1}{n_k} \sum_{i \in I_k} \frac{b(i) - a(i)}{max(a(i), b(i))} \right). \quad (2)$$

In addition to direct comparison, we also applied the load profiling result to infer the building operation schedule to demonstrate how the framework can assist in building energy modeling. While the result

can be used for other various applications, schedule inference was selected for this study because it is straightforward to understand and to benchmark. The inferred schedules were constructed for buildings according to eight date types: Weekdays, Saturdays and Sundays during semester; Weekdays, Saturdays and Sundays during vacation; holidays and days before holidays. The categorization is based on the analysis of the clustering results, which will be illustrated later. To automate this process, each sample was labeled with the date type and the date type of major samples in a cluster was used to represent the cluster. The centroids are taken as the schedule of the corresponding date types. The inferred schedules were then transformed to electricity load prediction and compared with the real data. The Mean Bias Error (MBE) and Coefficient of Variation of the Root Mean Square Error (CVRMSE) of the prediction were calculated for the buildings.

Results

Visual comparison

By investigating the identified load profiles and the corresponding days of each cluster, the clustering results were subjectively evaluated. Considering how informative and accurate the clusters were, the best methods for each building were selected. The proposed framework identified many interesting profiles and was more robust and intelligent in uncovering the building energy behavior from energy consumption data. Among the 50 buildings, the proposed framework gave the best clustering result for 37 buildings, K-means and DBSCAN respectively gave the best for 4 buildings and DTW gave the best for 1 building. There are another 3 buildings where the proposed framework gave the best, but the results still had some issues. There was one building that was too noisy for all the methods to work. Figure 3 shows the 5 methods' clustering result of 8 exemplary buildings. We used this figure to thoroughly elaborate all typical situations. The 2-dimension scatter plots were not shown here due to the space limitation.

Building 1 to 4 stand for the 4 typical types of the 37 buildings where the proposed framework outperformed the others:

- The first type included 11 buildings that operated differently during different periods in the year. In this example, the proposed framework discovered that the building weekday operation ended one hour earlier before August (cluster 2) than after August (cluster 1). While the other methods identified one or two clusters from the weekdays, none of them was able to well distinguish them.
- The second type accounted for the most buildings (19), where only the proposed framework was able to correctly identify the different profiles on different weekdays. From the image, it is obvious that profiles are different on Saturday (cluster 1) and Sunday (cluster 3). However, all other methods failed to separate them.
- 3 buildings from the third type contained a special profile happening only Feb. 18 and Dec. 24, when the operation was normal in the morning but off in the afternoon. These two days were one day before Chinese Spring Festival and Christmas. Note that there were other buildings following this pattern, but were not detected.
- There were other 4 buildings where the operation was affected by the academic semester. For example, building 4 consumed lower in the afternoon during vacation (cluster 4) than during semester (cluster 3), which was only recognized by the proposed framework. In this example, the proposed framework also had the advantage over the others of separating the Saturday and Sunday profiles clearly. These buildings with multiple advantages were categorized based on the dominant profiles.

The proposed framework also gave the most informative result to 3 other buildings, but they had issues of too similar clusters not merged. Take building #5 for example, Saturdays (cluster 4) and Sundays (cluster 5) were well separated from the weekdays, but two small clusters of normal weekdays that are supposed to be merged also stood out.

For the 8 buildings where K-means (e.g. building #7) or DBSCAN (e.g. building #6) gave the best result, the proposed framework identified too many meaningless clusters. Most of the residential buildings belonged to this category. Unlike all the other buildings, building #8 were very noisy and no meaningful cluster was identified either by visual observation or by the clustering methods. It was also noticed that the shape-based methods (DTW and K-shape) frequently gave bad unexpected results.

Table 1: Percentage of buildings where the proposed framework got higher score than other methods.

Sethods	Log(CH) (%)	Silhouette (%)
K-means	14.6	25
DBSCAN	79.2	77
DTW	22.9	35.4
K-shape	97.9	87.5

Benchmarking by Clustering Validation Index

Comparison between the proposed framework and the other 4 methods over the CH and Silhouette score was summarized in table 1. It can be seen from the table that the two CVIs basically gave the same result: Both K-means based methods got more higher scores than the proposed framework, the framework got more higher scores than DBSCAN, and K-shape got the lowest score. Comparing these with the visual evaluation results, it can be told that the visual

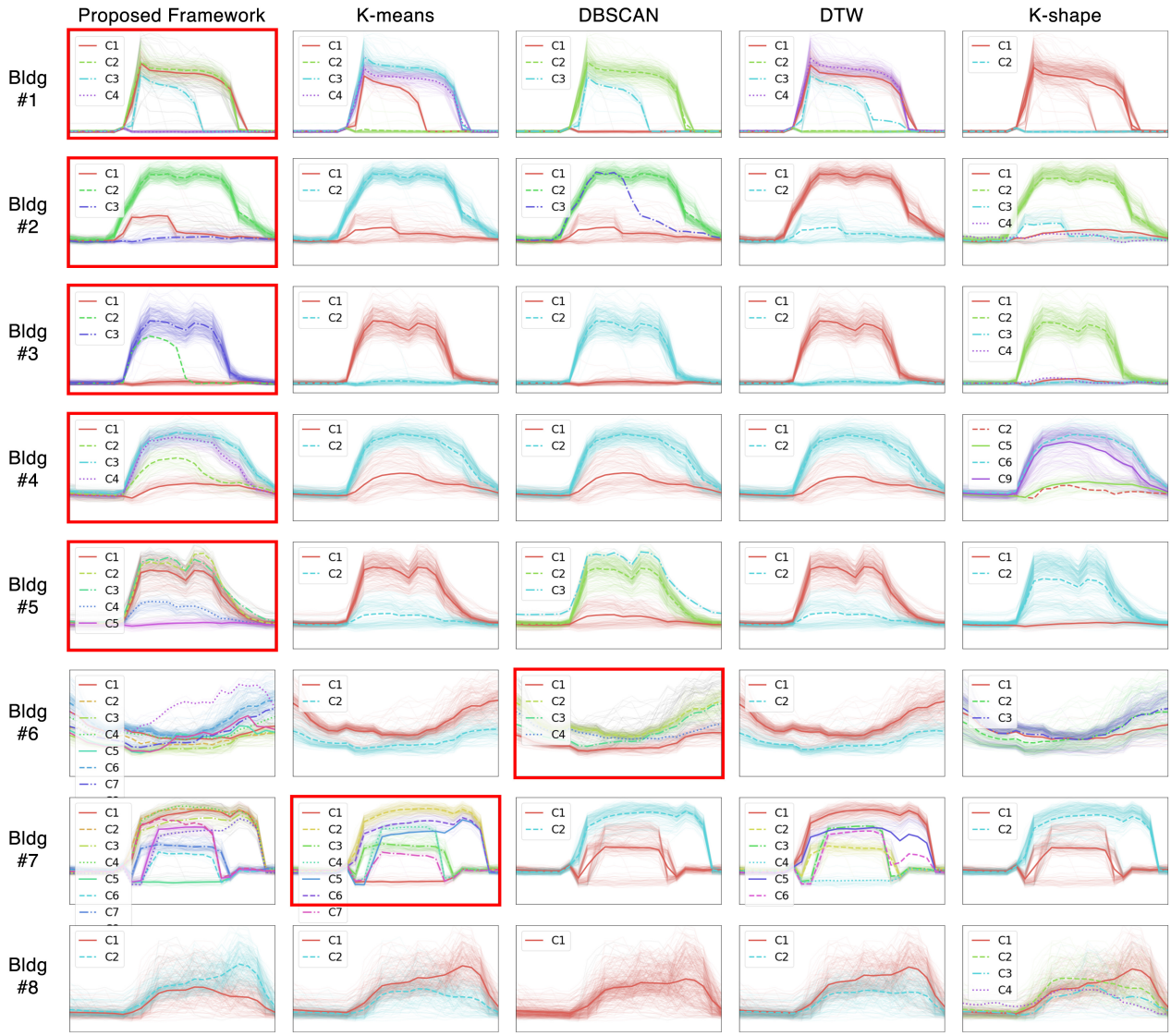

Figure 3: Clustering results of different methods on 8 exemplary buildings (best cases highlighted in red boxes).

observation did not match the CVI evaluation. The proposed framework got lower scores than K-means and DTW in most of buildings where it found finer clusters. Meanwhile, the proposed got higher scores than DBSCAN in the several buildings where DBSCAN gave the best clustering result. The reason of this paradox will be discussed later.

Benchmarking by load prediction

Comparison over the accuracy of building electricity load prediction was visualized in figure 4. Almost all MBE are close to 0 and therefore the logarithmic values are plotted for better visualization. In addition to the diagonal line for comparison, vertical and horizontal lines are also drawn in the CVRMSE plots for the 30% ASHRAE standard (ASHRAE, 2014). Since some clustering results were too trivial to extract operation pattern, several buildings were excluded for this benchmarking. For prediction errors, Smaller values mean better predictions. Thus, the red points lying in the lower half represent the buildings where the proposed framework gave better predictions. The percentage of red points out of the total number of points is reported for each plot. For easier comparison, the 37 buildings where the proposed framework was visually observed to be better are plotted as circles, and the others are plotted as crosses.

Since the centroids of each cluster were extracted as schedule, the positive and negative errors canceled each other when summed up. Therefore, MBEs of all the buildings were very small, close and not useful for further analysis. Meanwhile, the comparison result of CVRMSE agreed with the visual comparison. Regarding the standard, the proposed framework gave acceptable results for all buildings. Except that K-shape was much worse than all the others, the proposed framework outperformed all the other three methods in around 75% buildings. Most of the buildings, where the proposed framework was not the best solution, lay in the upper half of the corresponding plots. However, it is also noticeable that the scale of improvement was usually not very large since the extra schedules identified by the framework were usually on some minor day types.

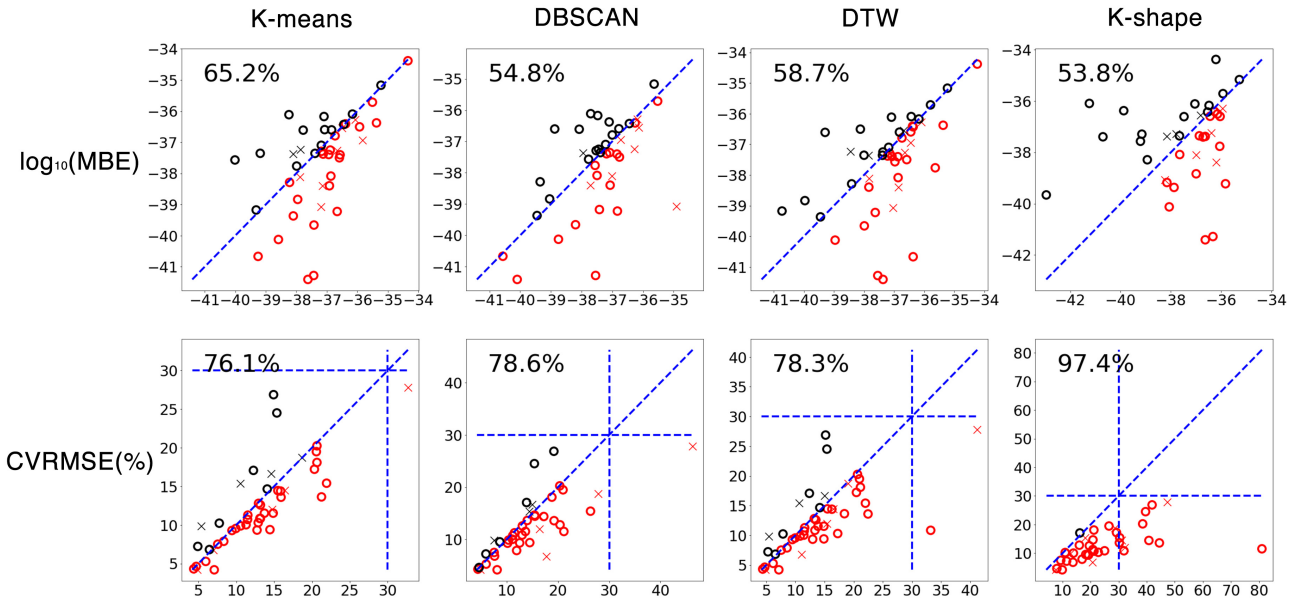

Figure 4: Comparison results of energy prediction accuracy (red/black: proposed got better/worse prediction; circle/cross: proposed got better/worse clustering).

Discussion

The effectiveness of the framework

According to the benchmarking result, the proposed framework outperformed the existing methods in different types of buildings and successfully captured the buildings' detailed energy behavior. The clustering result given by conventional methods like K-means might seem "acceptable", but there truly was information missed in comparison with that of the proposed framework. This knowledge, only discovered by the proposed framework, is valuable for further application. For example, a more precise and detailed building operation schedule helps with the building benchmarking and can lead to more accurate building energy simulation.

Why the proposed framework was more informative and robust has been elaborated before in the second section. However, while the framework solved the original problems of K-means and DBSCAN, it also brought a new potential issue of too many clusters identified. Among the 13 buildings where the framework failed, the most typical type was residential buildings. Like building #6 in figure 3, in addition to two major patterns respectively for semesters and vacations, residential buildings had many minor patterns caused by the occupants' variant behavior. In this situation, all the minor patterns were recognized as clusters at the second step and not merged at the final step, resulting in too many uninterpretable final profiles.

Thus, the high-resolution profiling worked well for most buildings but was too sensitive for buildings with too diversified behavior. Other than the residential buildings, undesired clusters were also generated for the rest of the 13 buildings. Taking building #5 in figure 3 as an example, the first three clusters are supposed to be merged at the last step but were not. This was caused by the drawbacks of the existing metrics, which we will discuss next.

The problem of metrics

Among many CVIs we tested, most were found not robust when applied to highly diversified datasets. For example, s_Dbw index was not applicable to high-dimensional data and Density Based Clustering Validation could easily get NaN. However, the CVI benchmarking results showed that even the selected two robust CVIs could not precisely describe the quality of the clustering result. As the example in figure 5, the proposed framework separated cluster 2 (Saturdays) and 3 (Sundays), which had clearly different profiles, but got lower scores than K-means. This is because these two clusters were close to each other and therefore lowered the scores. According to equation 1 and 2, both indices include elements representing the distance between clusters in the numerators to penalize too many clusters generated. However, the fact is some close and somewhat "similar" samples are actually different and should be separated. In other words, the distance is the only factor that defines clusters according to existing CVIs, while factors like density and continuity should also be considered. Fortunately, though the CH index cannot well evaluate the final clustering result, it was still reliable for tuning the parameters at each step in the framework. Take the case in figure 5 for example, the CH score reached 3577 after K-means. Then DBSCAN identified smaller and closer clusters within the preliminary clusters so that the maximum value of overall CH score was limited to 1935 before the last step. In this way, the framework took advantage of the CVI's capability of finding locally optimal solution and narrowed the range of CVI over the whole process so that the CH index's bias on close clusters was eliminated.

This is one of the reasons why the greedy algorithm was applied instead of global optimization.

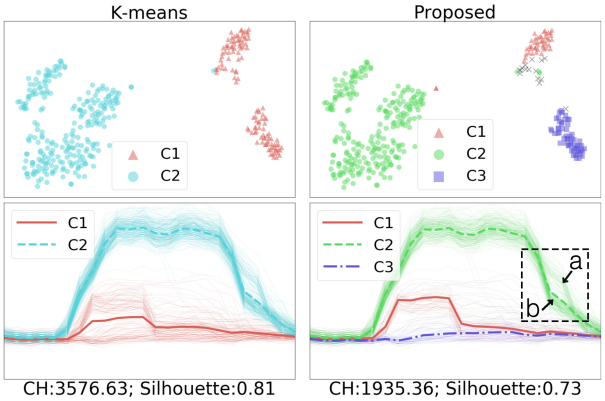

Figure 5: An example of failed CVI evaluation (Proposed got better result but lower scores).

Since no dimension reduction method was applied due to the potential information loss, the 24-dimensional data became very sparse in the space. This led to the problem of "Curse of Dimensionality". With 24 variables in each sample, the influence of stochastic error was essentially enhanced, sometimes overwhelming the true difference between samples. An example is highlighted by the black dashed rectangle in figure 5. Weekdays after October (a) consumed more than before (b) for 2 hours in the evening, but couldn't distinguish themselves from the others. The reason is that this 2 hours' difference, though obvious by observation, was not substantial in comparison with the variance of other hours when computing the Euclidean distance. Methods like DTW were proposed to overcome this kind of problems (Sakoe and Chiba, 1978) but unfortunately is not suitable here because the difference in scale and trend matters. PCC was used as a different metrics at the last step to complement for this weakness but still suffered similar problems, causing those undesired situations of similar final clusters not merged. Thus, applying a better and more suitable similarity metrics will further improve the framework's performance. Alternatively, different weights can be assigned to different hours to magnify the difference in critical hours, but the weighting strategy should be carefully designed to make it adaptive to different situations.

On further applications

The load prediction benchmarking result showed that the improvement of load profiling accuracy can easily result in improvement of further load prediction. Load prediction was implemented in a simple and scalable way to demonstrate the effectiveness. Two items to note here: 1) The prediction error was mainly caused by the lack of information. Since date type was the only input of prediction, the prediction accuracy can be improved by integrating more parameters like outdoor temperature; 2) To avoid overfitting, the approach of schedule generation cannot be reversed, i.e. to directly use the mean profiles of the 8 date types.

It is straightforward that finer profiles benefit energy prediction, especially when it comes to higher granularity. For other applications such as demand side management, the profiling results give more solid answer to questions including but not limited to "when will the peak load happen on certain date types" and "how much percentage of load can be shifted to nighttime". However, this study mainly focused on the profiling algorithm, questions like "how this information will affect the design strategy" and "how many extra saving can be achieved" remained unexplored. Also, note that the profiling resolution is not always the higher the better, but should adapt to the ultimate objective. For instance, if the profiles are used to classify the buildings or customers, too detailed profiles would lead to too large variance among buildings or customers and therefore inhibit the subsequent classification.

Conclusion

In this paper, we proposed a scalable unsupervised framework that can achieve high-resolution building load profiling. By applying K-means and DBSCAN complementarily, the framework is able to accurately identify the typical energy usage profiles for different types of buildings without any extra tuning. With the dataset of 50 buildings, the proposed framework has been proved to solve the problems of currently popular methods and outperform the existing methods. The framework provides a better approach to understanding how a building is operated. Also, the result is useful for further applications including but not limited to schedule inference for building energy modeling.

With a thorough comparison between the proposed framework and the existing methods, we revealed the reasons behind the performance improvement. Through further discussion, we identified three directions to extend this study: 1) A robust and more informative CVI is required to quantitatively evaluate the load profiling result; 2) A new similarity metric or a delicate weighting strategy is to be designed for time series clustering; 3) How this high-resolution profiling result can help tasks other than energy prediction is to be investigated.

References

ASHRAE (2014). Ashrae guideline 14–2014, measurement of energy, demand, and water savings.

Desgraupes, B. (2013). Clustering indices.

Eguaras-Martínez, M., M. Vidaurre-Arbizu, and C. Martín-Gómez (2014). Simulation and evaluation of building information modeling in a real pilot site. *Applied Energy 114*, 475–484.

Ester, M., H.-P. Kriegel, J. Sander, and X. Xu

(1996). Density-based spatial clustering of applications with noise. In *Int. Conf. Knowledge Discovery and Data Mining*, Volume 240.

Fan, C., F. Xiao, Z. Li, and J. Wang (2018). Unsupervised data analytics in mining big building operational data for energy efficiency enhancement: a review. *Energy and Buildings 159*, 296–308.

Fan, C., F. Xiao, and C. Yan (2015). A framework for knowledge discovery in massive building automation data and its application in building diagnostics. *Automation in Construction 50*, 81–90.

Green, R., I. Staffell, and N. Vasilakos (2014). Divide and conquer? k-means clustering of demand data allows rapid and accurate simulations of the british electricity system. *IEEE Transactions on Engineering Management 61*(2), 251–260.

Habib, U. and G. Zucker (2015). Finding the different patterns in buildings data using bag of words representation with clustering. In *Frontiers of Information Technology (FIT), 2015 13th International Conference on*, pp. 303–308. IEEE.

Hsu, D. (2015). Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data. *Applied energy 160*, 153–163.

Jalori, S. and P. T Agami Reddy PhD (2015a). A new clustering method to identify outliers and diurnal schedules from building energy interval data. *ASHRAE Transactions 121*, 33.

Jalori, S. and P. T Agami Reddy PhD (2015b). A unified inverse modeling framework for whole-building energy interval data: daily and hourly baseline modeling and short-term load forecasting. *ASHRAE Transactions 121*, 156.

Li, K., Z. Ma, D. Robinson, and J. Ma (2018). Identification of typical building daily electricity usage profiles using gaussian mixture model-based clustering and hierarchical clustering. *Applied Energy 231*, 331–342.

Liu, Y., Z. Li, H. Xiong, X. Gao, and J. Wu (2010). Understanding of internal clustering validation measures. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, pp. 911–916. IEEE.

Maulik, U. and S. Bandyopadhyay (2002). Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*(12), 1650–1654.

McLoughlin, F., A. Duffy, and M. Conlon (2015). A clustering approach to domestic electricity load profile characterisation using smart metering data. *Applied energy 141*, 190–199.

Miller, C., Z. Nagy, and A. Schlueter (2015). Automated daily pattern filtering of measured building performance data. *Automation in Construction 49*, 1–17.

Miller, C., Z. Nagy, and A. Schlueter (2018). A review of unsupervised statistical learning and visual analytics techniques applied to performance analysis of non-residential buildings. *Renewable and Sustainable Energy Reviews 81*, 1365–1377.

Panapakidis, I. P., T. A. Papadopoulos, G. C. Christoforidis, and G. K. Papagiannis (2014). Pattern recognition algorithms for electricity load curve analysis of buildings. *Energy and Buildings 73*, 137–145.

Paparrizos, J. and L. Gravano (2015). k-shape: Efficient and accurate clustering of time series. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, pp. 1855–1870. ACM.

Sakoe, H. and S. Chiba (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing 26*(1), 43–49.

Shahzadeh, A., A. Khosravi, and S. Nahavandi (2015). Improving load forecast accuracy by clustering consumers using smart meter data. In *Neural Networks (IJCNN), 2015 International Joint Conference on*, pp. 1–7. IEEE.

Tang, F., A. Kusiak, and X. Wei (2014). Modeling and short-term prediction of hvac system with a clustering algorithm. *Energy and Buildings 82*, 310–321.

Tsekouras, G. J., N. D. Hatziargyriou, and E. N. Dialynas (2007). Two-stage pattern recognition of load curves for classification of electricity customers. *IEEE Transactions on Power Systems 22*(3), 1120–1128.

Verleysen, M. and D. François (2005). The curse of dimensionality in data mining and time series prediction. In *International Work-Conference on Artificial Neural Networks*, pp. 758–770. Springer.

Wei, Y., X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, and X. Zhao (2018). A review of data-driven approaches for prediction and classification of building energy consumption. *Renewable and Sustainable Energy Reviews 82*, 1027–1047.

Yang, J., C. Ning, C. Deb, F. Zhang, D. Cheong, S. E. Lee, C. Sekhar, and K. W. Tham (2017). k-shape clustering algorithm for building energy usage patterns analysis and forecasting model accuracy improvement. *Energy and Buildings 146*, 27–37.