Clustering Energy Load Curves of Commercial Buildings:
Using Typical Load Curves for Building Type Clustering

Juan Tzoc
Carnegie Mellon University
jtzoc@andrew.cmu.edu

Jiaming Xiao
Carnegie Mellon University
jiamingx@andrew.cmu.edu

ABSTRACT
Given the influx of inexpensive, advanced smart meters available on the market, there is significant potential to apply their detailed collected data to better understand and optimize commercial building energy usage. The load curves of commercial buildings, collected from smart meters, can be clustered to create typical load curves of buildings with similar attributes. Facility managers can use these typical load curves to benchmark their building, to determine their building's performance compared to buildings with similar attributes, such as location, type, and layout. This comparison will allow facility managers to determine whether similar buildings are implementing energy-efficient or energy management techniques that can be replicated to improve their own building's energy usage. In this paper, several different types of commercial buildings from 4 TMY3 locations in Pennsylvania were used to look for similar patterns among commercial buildings with similar characteristics, specifically building type. The K-Means Clustering algorithm is deployed to cluster these buildings' load curves to create typical load profiles. Results indicate that typical load profiles can be created for commercial buildings when using clustering techniques that are reflective of building type and other similar building characteristics.

Keywords
Benchmarking; K-Means Clustering; Smart Meters;

1. INTRODUCTION
It is estimated that 40% of the annual total energy consumed in the United States is contributed by the building sector [6]. With the recent increase of inexpensive, advanced smart meters hitting the market, "at high resolutions in the order of minutes to seconds" [2], there is significant potential to "understanding and optimizing [both residential and commercial] building energy usage" creating opportunities to reduce total energy consumption in the United States [2].

Having the ability to track gas and electricity consumption of commercial buildings at finer granularity provides ample opportunity for facility managers to generate more detailed benchmarks against similar buildings to evaluate key differences and "potential energy-efficiency improvements and optimizations" [2]. Facility managers can effectively compare their building's energy usage to similar buildings which can provide what portions of consumption are due to permanent building characteristics (i.e. climate region, building layout, building type, etc.) versus occupant and energy management behaviors [4]. Providing these portions can help facility managers determine if behavior can be changed to reduce energy consumption [4].

Our group reviewed the following literature to determine an effective way of benchmarking a commercial building against similar buildings using energy load curves to establish whether there were any potential energy-efficiency improvements:

In [2], individual homes are decomposed "into different groups based on the characteristics of their load profiles" [2], which reflect daily routines and patterns.

In [4], an interface is created for users to enter electricity and gas usage to determine how their energy consumption compares to similar homes predicted by a model generated from a data set of monthly utility bills of approximately 6,500 buildings over several years.

In [1], three clustering techniques, including K-means, were used to classify electricity customers based on time-of-day and consumption usage characteristics for determining abnormalities in consumption patterns, like faulty metering and illegal connections, to trigger alerts.

2. PROPOSED APPROACH
The goal of our project is to cluster anonymously labeled load curves, using K-means clustering, to see if similar building characteristics, like building type, can be extracted from these clusters. From these clusters, our group can generate typical load profiles for similar building characteristics within each cluster, which can be used as comparative models. These models can then be compared against a building, with shared attributes, to determine similarities and differences in energy usage. Comparing the model against a building with similar attributes can indicate whether there are issues of energy-efficiency or negative energy management behavior for that particular building or if this building has implemented energy saving features or behaviors that similar buildings can also implement.

3. DATASET
Our group utilized commercial building hourly load profiles, for all TMY3 (typical meteorological year) locations in the United States dataset from [3] for classifying building characteristics from annual load curves. The dataset contains hourly load profile data for 16 commercial building types, as well as residential buildings, for each of these TMY3 locations [3]. For each location and building type there is hourly data for facility electricity, facility gas, cooling electricity, heating electricity, heating gas, and lighting electricity consumption [3].

These 16 commercial building types are based off the DOE commercial reference building models [3], which are commercial building benchmark models that reflect different types of commercial buildings in various climate zones across the United States [5]. Additionally, these 16 commercial building types "represent approximately 70% of the commercial buildings in the U.S." and include the following building types: offices, warehouses, schools, hotels, hospitals, and supermarkets [4, 6].

For our analysis, 4 TMY3 locations were chosen in the state of Pennsylvania, specifically, Bradford, Butler, Philadelphia, and

Pittsburgh. To further simplify our analysis, only facility gas and electricity consumption were used to determine building characteristics, such as building type and location based solely on the annual load curves of a building. Our group chose to work with a smaller set of this dataset as a platform for further work and improvements.

4. RESULTS

By manipulating the aforementioned datasets, hourly electricity and gas load profiles of hospital, hotel, office and supermarket buildings in Bradford, Butler, Pittsburgh and Philadelphia were generated. Box plots and stem plots are deployed to visualize these load profiles to highlight some interesting patterns. To find these patterns all the features in the data sets were used. However, the most applicable data for comparison use are the total electricity and gas load profiles, as they proved to be the best attributes in the dataset to analyze building types and other similar characteristics.

4.1 Analysis of Energy Consumption Based on Building Type

For commercial buildings, functionality may be a significant factor that determines their energy consumption profile. To observe load profile patterns among different types of buildings, our group deployed box plots to visualize the load profile of each hospital, hotel, office and supermarket in the four Pennsylvania sites. Results are shown in Figure1 and Figure 2.

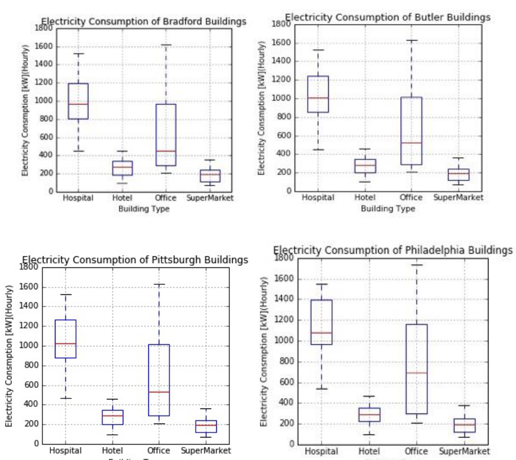


Figure 1: Electricity Hourly Consumption

From Figure 1, it can be observed that the electricity load profiles of those four types of buildings are slightly different among the four places sampled. Compared to other building types, electricity load profile of the hospital is relatively higher. This phenomenon can be explained by the fact that there are a lot of patients who live in hospital 24 hours each day. To ensure the health and safety of those patients, facilities and medical devices in the hospital need to keep working day and night. This may lead to higher average hourly electricity consumption. It can also be observed that variance of the office's electricity load profile is higher than other building types. This is possibly because occasional or seasonal overtime work can create variations in daily or weekly electricity consumption. By observing the four subplots in Figure 1, it can be seen that the electricity load profile differences among the types of buildings show the same pattern in the four sampled places.

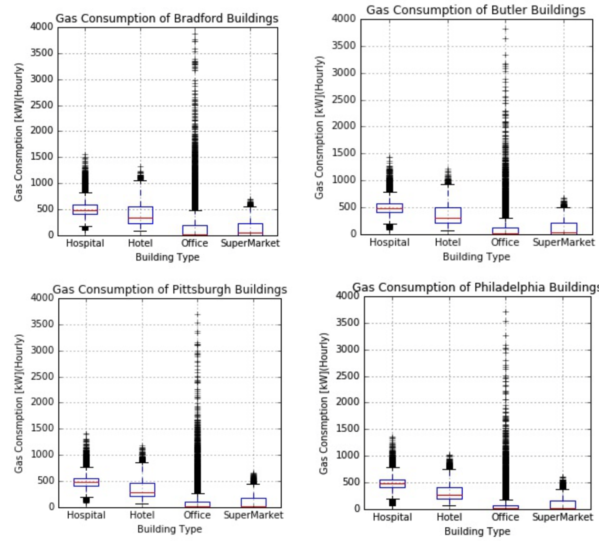


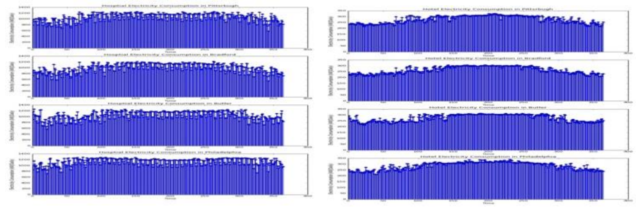
Figure 2: Gas Hourly Consumption

In Figure 2, it can be seen that the gas load profile of the hospital shows the same pattern as the electricity load profile, which is higher than other building types. But this difference is not as large as in the electricity load profile. This may be because most of the devices in hospitals are electricity driven, which means gas consumption of the hospital is almost at the same level as other buildings. The variance of hourly gas consumption in the office is extremely high. This can be attributed to some extreme local site conditions that can't be evidently proved or detected with the datasets in hand. More data are needed to explain the outliers within office hourly gas consumption. Like Figure 1, it shows that gas load profile difference among those four building types in the four sampled places share the same pattern, which means hourly electricity consumption and hourly gas consumption can be deployed for the purpose of building type clustering.

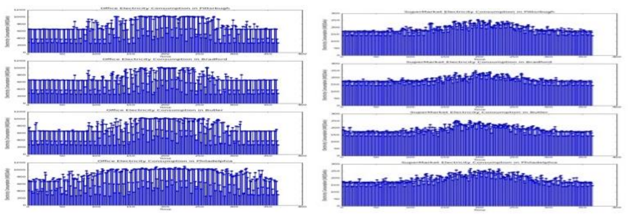
4.2 Load Curve Analysis for Different Building Types

It was observed that the electricity and gas load profile for each type of building in Bradford, Butler, Pittsburgh and Philadelphia have similar patterns. At this time, our group will observe load curves of a building type, such as a hospital, at different locations. Stem plots are built to show the load curves. Results are shown from Figure 3 to Figure 4.

From the stem plots, it can be seen that electricity load profile and gas load profile are complementary, which is consistent with the energy consumption pattern in Pennsylvania, in which most buildings are cooled by electricity and heated by gas. Moreover, energy consumption for each of the four buildings at different locations show the same pattern, again indicating that electricity and gas consumption are the attributes that can be used to predict building type.

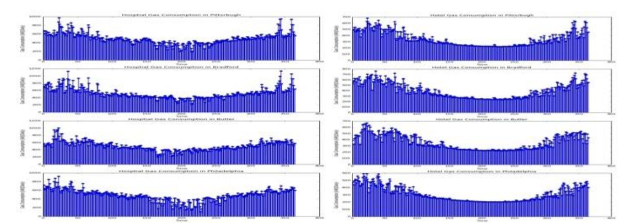


Hospital (a) Hotel (b)

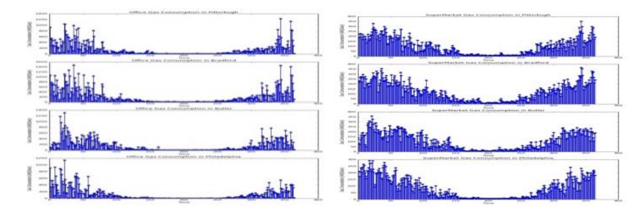


Office (c) Super Market (d)

Figure 3: Daily Electricity Consumption



Hospital (a) Hotel (b)



Office (c) Super Market (d)

Figure 4: Daily Gas Consumption

4.3 K-means Clustering for Building Types

After analyzing the electricity and gas load profile, our group found that in each of the sampled places, different building types show distinguished load profiles. Additionally, each type of building in the four sampled places show similar load profiles. It is suggested that electricity and gas consumptions are helpful for clustering. Thus, a k-means clustering algorithm is deployed to explore building type clustering based on electricity and gas load profile. Result is shown in Figure 5.

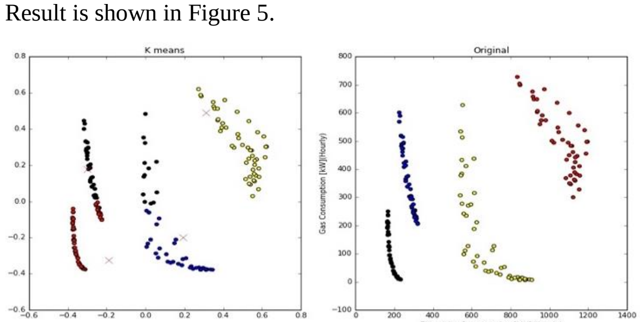


Figure 5: K means clustering

For the right (reference) scatterplot, in Figure 5, an attribute is added labeling each sample. The attribute value ranges from 0 to 3. 0 represents hospital, 1 represents hotel, 2 represents office and 3 represents supermarket. In the right scatterplot, in Figure 5, it can be observed that the right corner cluster represents the group of hospitals since samples in this cluster shows a higher gas load profile and a higher electricity load profile. The cluster in the middle area represents the group of offices since the electricity load profile in this group are the second largest ranges. The most left corner cluster represents the supermarket which shows the lowest range of both electricity and gas load profiles. The cluster situated between supermarket and office cluster is the hotel cluster.

By observing the k-means result, in the Figure 5, it can be found that it shows almost the same clustering pattern as the reference plot, except for a part of overlap between the supermarket cluster, which is at the left corner, and the hotel cluster. Figure 5 also indicates that electricity energy and gas load profiles are recommended attributes to make the building type clustering.

5. DISCUSSION

From our results, it was found that electricity and gas load profiles of buildings at some level reflect the building's functionality, which means they can be used to create typical building load profiles. Additionally, our group found that differences in electricity and gas load profiles among different building types show the same pattern at the four sampled places. This may be because those places share similar location features. More samples are needed to prove that the pattern is not a coincidence. The k-means clustering results indicate that electricity and gas load profiles of buildings are good attributes to cluster buildings with different building types.

6. FUTURE WORK

In this paper, our group analyzed electricity and gas load profiles of different types of buildings at four places (Bradford, Butler, Pittsburgh and Philadelphia). The sample number is limited. In the future more samples should be added, to ensure that the patterns we observed in this paper are not coincidence. Our group will try to use different classification algorithms to classify building type based on building electricity and gas consumption, to discern which algorithm is better for building type classification and typical load curves.

7. REFERENCES

[1] H. Nizar, Z. Dong., and J. Zhao. Load profiling and data mining techniques in electricity deregulated market. In IEEE Power Engineering Society General Meeting, 2006. DOI= http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1709335

[2] Iyengar, Srinivasan, et al. Analyzing Energy usage on a City-scale Using Utility Smart Meters. University of Massachusetts Amherst, n.d. Print. 1 Nov. 2016.

[3] Wilson, Eric. Office of Energy Efficiency & Renewable Energy (EERE). Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States. U.S. Department of Energy Open Catalog. Web. Nov 2016. DOI= http://en.openei.org/doe-opendata/dataset/commercial-and-residential-hourly-load-profileS-for-all-tmy3-locations-in-the-united-states

[4] Z. Kolter and J. Ferreira. A large-scale study on predicting and contextualizing building energy usage. In AAAI, 2011. Web. Dec. 2016. DOI=

http://www.zicokolter.com/wpcontent/uploads/2015/10/kolter-aaai11.pdf

[5] Office of Energy Efficiency & Renewable Energy (EERE). Commercial Reference Buildings. U.S. Department of Energy. Web. Dec. 8 2016. DOI= http://energy.gov/eere/buildings/commercial-reference-buildings

[6] U.S. Energy Information Administration. Web. 2016. https://www.eia.gov/