

Article

Classification of Household Appliance Operation Cycles: A Case-Study Approach

Zeyu Wang ^{1,*} and Ravi S. Srinivasan ²

¹ M.E. Rinker, Sr. School of Construction Management, 324 Rinker Hall, University of Florida, Gainesville, FL 32611, USA

² M.E. Rinker, Sr. School of Construction Management, 316 Rinker Hall, University of Florida, Gainesville, FL 32611, USA; E-Mail: sravi@ufl.edu

* Author to whom correspondence should be addressed; E-Mail: wangzeyu@ufl.edu; Tel.: +1-213-400-7098.

Academic Editor: Nyuk Wong

Received: 8 July 2015 / Accepted: 17 September 2015 / Published: 22 September 2015

Abstract: In recent years, a new generation of power grid system, referred to as the Smart Grid, with an aim of managing electricity demand in a sustainable, reliable, and economical manner has emerged. With greater knowledge of operational characteristics of individual appliances, necessary automation control strategies can be developed in the Smart Grid to operate appliances in an efficient manner. This paper provides a way of classifying different operational cycles of a household appliance by introducing an unsupervised learning algorithm called k-means clustering. An intrinsic method known as silhouette coefficient was used to measure the classification quality. An identification process is also discussed in this paper to help users identify the operation mode each types of operation cycle stands for. A case study using a typical household refrigerator is presented to validate the proposed method. Results show that the proposed the classification and identification method can partition and identify different operation cycles adequately. Classification of operation cycles for such appliances is beneficial for Smart Grid as it provides a clear and convincing understanding of the operation modes for effective power management.

Keywords: smart grid; refrigerator operation cycles; k-means clustering; unsupervised learning

1. Introduction

For a century, utility companies have used conventional power grid system to provide electricity. Lack of techniques to automatically monitor the grid system and transfer useful data has forced utility companies to send workers out to gather data needed for optimal management of power grid systems. For example, these utility workers perform multiple functions on-site including reading meters, looking for faulty equipment, and measuring voltage. In such a conventional and, potentially outdated power grid system, there is little chance to optimally manage supply and demand effectively. In recent years, a new generation of power grid system, referred to as the Smart Grid, with an aim of managing electricity demand in a sustainable, reliable, and economical manner has emerged [1]. The Smart Grids European Technology Platform considers Smart Grid as an electricity network that can intelligently integrate the actions of all the connected users, generators, consumers, and those that do both, in order to efficiently deliver sustainable economic and secure electricity supply [2]. The Smart Grid is made possible by two-way communication technology and automation that has been used for decades in other industries. The two-way communication technology, which can transmit and receive digital signals, enhances the information transfer between each devices associated with the grid, e.g., substations, transformers, switches, consumers, and central control center. The automation makes it possible to adjust and control, potentially, each individual device or a collection of devices from a central location. Energy efficiency is achieved by employing appropriate power control algorithms. In general, Smart Grid is an intelligent power system that integrates advanced control, communications, and sensing technologies into the power grid.

Smart Grid contributes significantly in peak demand management as the trend of power demand can be monitored and the demand can be shifted from “on-peak” to “off-peak” periods by employing Demand Response [3]. However, applying Smart Grid for smart control of each consumer’s, for example, operation status recognition, operation control optimization, fault detection, etc., is challenging owing to uncertainty in usage conditions of household [4]. Nevertheless, using bi-directional data transfer enabled via Smart Grid and with greater knowledge of operational characteristics of individual appliances, necessary automation control strategies can be developed to operate appliances in an efficient. In order to provide specific power utilization information, effective recognition of the types of appliances have been considered by many researchers [5–7]. These research works demonstrate that multiple appliances can be distinguished based on certain classification algorithms. It is to be noted that only a handful of research work [8] has been done to study the classification of multiple operation modes of appliances. Some of the household appliances, such as a lamp or television, have single or limited operation modes, while others such as a refrigerator or clothes washer, have multiple operation modes due to their settings. In general, an equipment operation cycle may have unique characteristics such as “run time”, “idle time”, “maximum power”, etc., based on the operation of the appliance. Classification of operation cycles for each appliances is beneficial for Smart Grid as it provides a clear and convincing understanding of the operation modes for effective power management. This paper discusses a classification method to classify multiple operation cycles for a household refrigerator.

In recent years, the research community has used the concept of load signatures for appliance load monitoring. The load signature corresponds to the specific electrical behavior of an individual appliance/piece of equipment when it is in operation [9]. A load signature could use multiple attributes

such as active power, reactive power, switching transient current, and waveform to describe the transient and steady-state behaviors of an appliance. In fact, as each appliance has a unique load signature, many research studies focused on using load signature to classify household appliances. Huang *et al.* developed a household appliance classification method by using appliance power signature and harmonic features [10]. Xu and Dong formulated an appliance signature vector which comprises of different load signatures to identify household appliances [11]. Chui *et al.* used four appliance features namely, active power, reactive power, total harmonic distortion, and maximum transient current for household appliance classification and identification [7].

Meanwhile, many algorithms have been applied to identify different load signature patterns. Srinivasan *et al.* proposed a neural network based approach for signature identification of electrical devices based on the current harmonics [6]. Popescu *et al.* characterized home appliance load signature from the perspective of Recurrence Plot Analysis [12]. Kim *et al.* used the k-nearest neighbor method to detect specific appliance electrical events [13]. Also, clustering as a main technique to partition data into groups has been used to classify building electricity customers [14], predict future building energy demand [15], and detect abnormal behaviors [16]. Chui *et al.* proposed an appliance signature identification solution using k-means clustering to identify different household appliances [7].

Although, previous research efforts have proven load signature as an effective mean to classify and identify different appliances, there is no research focusing on household appliance operational pattern classification and identification. The classification and identification of these operational patterns could help users to understand their operation modes, estimate their energy usage, and diagnose the health of their appliances. This study fills this gap by introducing k-means clustering technique which is one of the most popular unsupervised learning algorithms that solve the clustering problem to classify the operational cycles of typical household appliances. The objective of the research work is to aid the operation mode recognition method of household appliances through classifying and recognizing different operation patterns. To validate the proposed classification method, a case study focusing on pattern classification of undisturbed household refrigerator operation cycles is conducted.

Household Refrigerator

The household refrigerator is selected for the case study as it is as one of the major household appliances which used 8% of U.S. residential electricity use in 2012 [17]. While refrigerator energy usage cannot be ignored, the use of refrigerants in refrigerator equipment also adds to the degradation of the environment, specifically the greenhouse gas emissions and ozone layer depletion [18]. Needless to say, smart control of household refrigerators is of great importance for achieving sustainability.

The operation of refrigerators is impacted by factors such as frost formation, door opening, and defrosting. Previous research has indicated that the energy use of a refrigerator can be dynamically modeled by considering the impacts of these factors [19]. However, the derivations of these factors are different, for example, frost formation and defrosting are determined by refrigerator's own factory settings and control strategy, while door openings are derived from user interference. Previous research has shown that occupant interference has a significant impact on refrigerator energy use, specifically owing to the frequency of door opening [20]. Extending these concepts, in our study, we broadly consider user interferences in two scenarios—disturbed and undisturbed. During the disturbed scenario,

user interference occurs and vice-versa. For this reason, this paper uses undisturbed cycles as they more persuasively represent typical refrigerator operation. Moreover, efficiency strategies to improve refrigerator health and performance may be implemented via Smart Grids during undisturbed periods. The other factors that impact energy use are the type of refrigerator door construction including the insulation technologies, and the refrigeration electrical components and control technologies. Thus, in this paper, we focus on the pattern recognition of undisturbed cycles in order to extract the household refrigerator's operation characteristics.

The paper is organized as follows: typical refrigerator operation cycles are introduced in Section 2. In Section 3, we present the basic idea of the proposed research methodologies including k-means clustering, attributes selection for clustering, selection of number of clusters, measurement of clustering quality, and identification of different clusters. Section 4 presents a case study to validate the proposed method. In Section 5, we discuss the conclusion and future works.

2. Typical Refrigerator Operation Cycles

Household refrigerators are cyclic appliances whose operations consist of repeating cycles. A typical refrigerator cycle includes both run time and idle time. While run time is when the refrigerator turns on to refrigerate or to defrost, idle time is when the refrigerator turns off and, therefore, no electricity is consumed. These operation cycles are further divided into two major groups: disturbed and undisturbed cycles, based on occupants inference such as door opening and loading/unloading food. Since operation cycles are impacted by occupants' usages which typically happen in daytime, the undisturbed cycles occur consecutively in a time period that happens between late night and early morning when there is little or no interference from occupants. Figure 1 shows a typical operation cycle which normally occurs without occupant's interference. Many alternative attributes which describe the characteristics of the operation cycle can be extracted based on Figure 1. These attributes include, but are not limited to startup time, operation time, shut down time, run time, idle time, total electricity use, average power, maximum power, and minimum power. In order to obtain satisfactory classification results, selecting appropriate attributes is crucial for this study. The selection of attributes is addressed in detail in Section 4.2—data preprocessing.

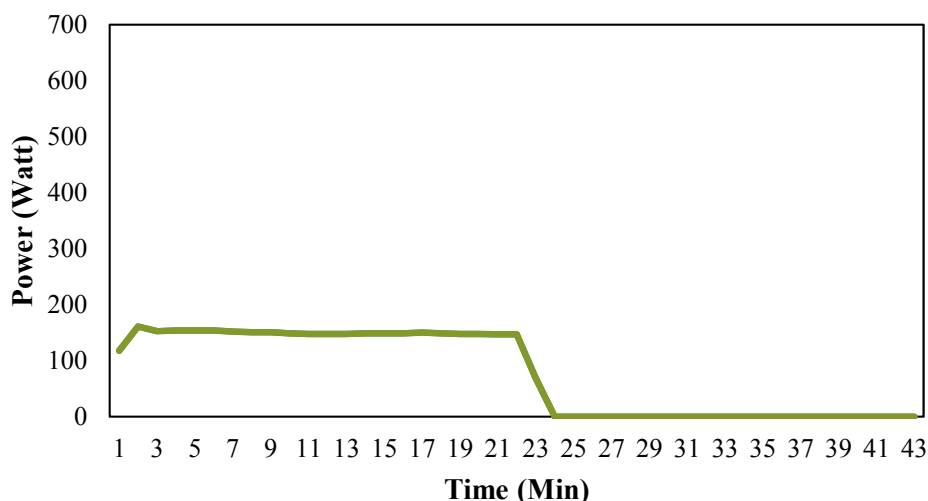


Figure 1. Typical household refrigerator operation cycle for 24 h period.

3. Methodology

The pattern classification strategy developed in this study comprises of four steps namely, data collection, data preprocessing, k-means clustering, and pattern recognition, Figure 2. In general, appliances energy use data were collected by using innovative monitoring meters. The operation cycle is identified by the turn-on and turn-off events of the appliances. For each operation cycle, appropriate attributes which are capable of describing the difference between different types of operation cycles should be selected. K-means clustering algorithm is applied to partition the data into several groups. As k-means clustering method requires pre-determination of correct number of clusters, an error function which can assist this determination in selecting optimum k is introduced. Because the ground truth is unknown for this study, an intrinsic method, *i.e.*, silhouette coefficient is used to measure the clustering result. Finally, we recognize and define the clustering results based on their characteristics such as average total energy use, average maximum power, sequential order, and frequency.

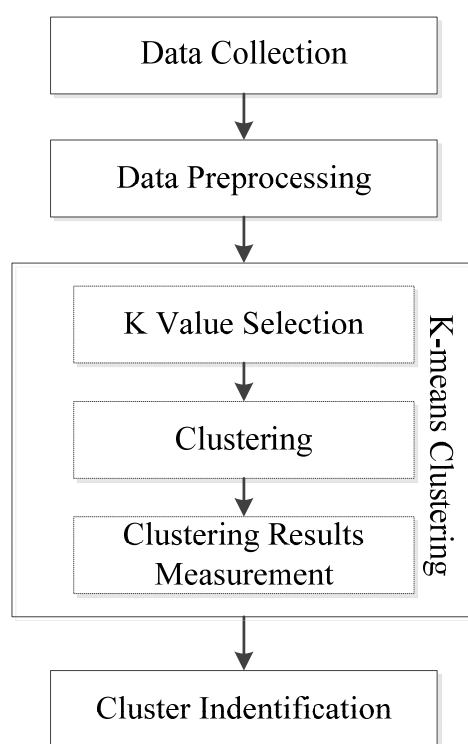


Figure 2. Flowchart for pattern recognition of different refrigerator operation cycles.

3.1. K-Means Clustering

Clustering is an unsupervised learning technique that groups a set of data into multiple clusters so that objects in the same cluster have high similarity, while they are very dissimilar to objects in other clusters [21]. In other words, clustering technique groups the data objects based on whether they are similar or not, and the similarity and dissimilarity are assessed based on the attributes which are used to describe the objects.

In this paper, we applied a centroid-based clustering algorithm, k-means clustering, for pattern recognition as it is efficient for large and multi-dimension datasets. The basic idea of k-means algorithm is to assign objects to the most similar cluster and update the center value of the cluster iteratively until

the assignment is stable, that is, the clusters formed in current round are the same as those forms in the previous round. An error function is introduced to verify the stability of the assignment and measure the partition quality. In k-means, a cluster is represented in terms of its centroid which is defined as the mean value of the objects within the cluster. The difference between each object in the cluster and the cluster centroid is measured by the Euclidean distance which is defined as follows:

$$d_{euclidean}(x, C) = \left[\sum_{i=1}^d (x_i - c_i)^2 \right]^{\frac{1}{2}}$$

where x is the object, C is the cluster centroid, d is the dimension of the data, x_i is the i th attribute value of the object, c_i is the i th attribute value of the cluster centroid.

3.2. Attributes Selection for Clustering

Since clustering determines the similarity and dissimilarity based on the attributes describing the objects, appropriate attributes which can help distinguish objects of different groups should be selected and extracted from the initial database. As noted previously, many attributes of the operation cycle can be considered for clustering. While in this paper, the refrigerator operation cycles are distinguished from the perspective of different working modes, for example, auto-defrost mode, general operating mode, and cooling mode. In general, different working modes of operation that may result in different operation time, power rate, waveform, *etc.* Therefore, certain attributes which are able to describe such difference should be selected in order to receive promising classification result. Notably, there is no uniform selection of attributes that can classify operation cycles of appliances of all kinds. Different types of r appliances could have different operational control strategies, working components, *etc.*, thus their differences between different working modes are hard to be described in a uniform manner. The selection of attributes highly depends on the type of appliance and the way it shows the difference between different operation modes. Preliminary studies and expert knowledge are needed to investigate appropriate attributes that can help in distinguishing different operation modes for a certain types of appliance.

3.3. Selection of Number of Clusters

The exact number of clusters is required to be pre-determined for k-means clustering method since appropriate number of clusters controls the proper granularity of cluster analysis [21]. The precise number of clusters should be able to build a good balance between accuracy and compressibility in cluster analysis. In this study, the right number of clusters is ambiguous since the dataset is unlabeled and the number of different refrigerator operation cycles is unknown. The determination of the correct number of clusters depends on the distribution's shape and scale in our dataset, as well as the clustering resolution required for this research work. An error function is introduced to help find the exact number of clusters for this study. An error function is defined as:

$$E = \sum_{i=1}^k \sum_{x \in C_i} d(x_i, C_i)$$

where E is the sum of the squared error for all objects in the dataset, x is the data point in each cluster, C is the central point of each cluster, i is the number of clusters, and d is Euclidean distance between each data point x and its central point C . The error function aims to make the resulting k clustering as compact and as separate as possible. In other words, k -means algorithm can be treated as an optimization problem which minimizes the error function. The process of selecting the exact number of clusters is described as follows:

Step 1: Selection of an alternative k ($k > 0$) values which are likely to be the right number of clusters based on the distribution of the operation cycles;

Step 2: Calculation of the total error value of the selected k based on the error function;

Step 3: Building a curve of total error with respect of k which shows its tendency;

Step 4: Finding an optimum k value based on the total error trends. Normally, the first or most significant turning point of the curve suggests the right number of clusters [21]; and

Step 5: Validating the selected k according to the distribution and the resolution required by the researcher.

3.4. Measurement of Clustering Quality

In this study, since the ground truth which is the number of different types of refrigerator operation cycles is unavailable, we measure the clustering quality by using the intrinsic methods which evaluate the goodness of a clustering by considering how well the clusters are separated. The silhouette coefficient which was first proposed by Peter J. Rousseeuw is a method to show how well each object lies within its cluster [22]. The silhouette coefficient is defined as:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

where i is an individual point in the dataset, $a(i)$ is the average distance of point i to all other objects in its cluster, $b(i)$ is the minimum average distance of point i to points in other clusters. The value of the silhouette coefficient is between -1 and 1 . As $a(i)$ measures how dissimilar i is to its own cluster, a smaller value means the cluster is more compact. The value of $b(i)$ implies the degree of difference between i and other clusters, thus the larger $b(i)$ is, the more separated i is from other clusters. Therefore, positive silhouette coefficient value means the cluster including i is compact and i is far from other clusters, while negative silhouette coefficient value means i is closer to the objects in another cluster than to the objects in its own cluster. Normally, for silhouette coefficient value, the bigger, the better, and value 1 is the extreme preferable condition.

3.5. Identification of Different Clusters

Identification of different clusters is important for users to understand their operation modes as the classification partitions cycles on a statistical basis, which is impossible to reveal the meaning behind each clusters. In this paper, the concept of identification of different clusters is to collect information which is able work as clues to link each cluster to the specific operation mode. Specifically, clusters are identified based on their characteristics such as volume, average power, average energy use, and sequential order, which are able to reflect the relation between clusters and operation mode.

4. Case Study

4.1. Data Collection

The experimental data used for method validation were obtained by monitoring a trio bottom freezer refrigerator with external filtered water dispenser manufactured in 2005. The experiment was conducted in an occupied two story residential house located in Lincoln, Nebraska. A nonintrusive monitoring system called eMonitor [23] was used to monitor minute-by-minute electricity use data at the circuit level. To obtain sufficient data, the monitoring period started from 16 January 2011 to 15 April 2011 with a total number of 83 calendar days. As the interval equals to one minute, a total number of 119,520 data points (1440×83) were collected for each circuit. The minute-by-minute energy use data of refrigerator allows the determination of the paired turn-on and turn-off events of the refrigerator. Each operating cycle can be identified from such turn-on and turn-off events. A total number of 2165 cycles have been identified from the monitored data.

4.2. Data Preprocessing

As shown in the Figure 3, a comparison of electricity use flows between different operation modes is conducted in order to find out appropriate attributes.

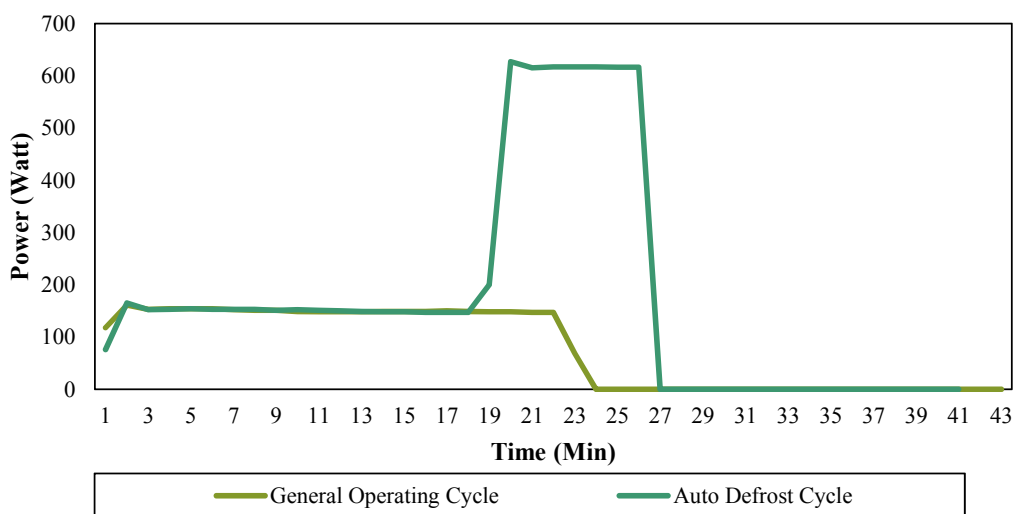


Figure 3. General operating cycle vs. auto defrost cycle.

The comparison indicates that auto defrost cycle has a significant higher power which is around 650 Watt. The run time and idle time are 27 min and 14 min respectively. It can be concluded that auto defrost cycle results in more energy use than general operating cycle by increasing run time, decreasing idle time, and switching to high level. As a result, two attributes namely, the total electricity use which reflects both the run time and average power, and the maximum power which describes the freezer working modes are selected as the attributes to describe the operation cycle.

4.3. Study of Refrigerator Hourly Energy Performance

An initial study focusing on refrigerator hourly energy performance was conducted in order to understand the daily variation tendency of refrigerator energy use. Figure 4 depicts the hourly average energy use tendency. From Figure 4, the fluctuation of trend line indicates that energy use is highly affected by occupants' interference. For example, the hourly energy use between 0:00 a.m. to 8:00 a.m. is approximately 30% less than all others. The difference indicates that during midnight when there is less or no interference from the occupants, the refrigerator consumes less energy than that of the daytime when occupants use refrigerator frequently. Another representative example is the peak load during meal time; three peaks are noticed in Figure 4 at time periods (represented in 24 h format): 8 to 9; 10 to 13; and 18 to 20 respectively. These three time periods are highly related to occupants' meal time. It is not difficult to speculate that occupants will use the refrigerator more frequently at meal time than any other time and as a result of that, more energy will be consumed.

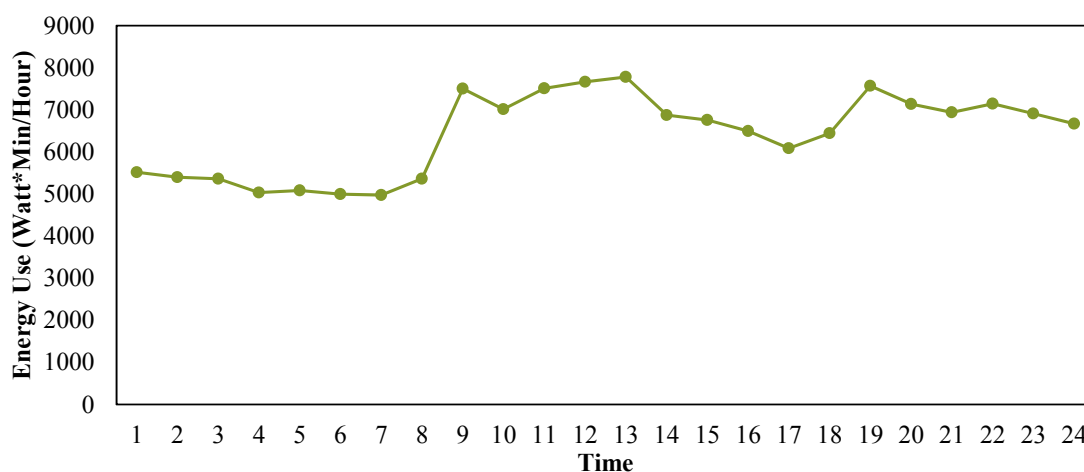


Figure 4. Hourly average energy use.

From this initial study, it can be noted that the energy efficiency can be evaluated from the time period when less or no occupants' interference occurred and refrigerator runs stably. Based on this, we used time period of operation cycles that run between 1:00 a.m. to 6:59 a.m. as the major source of undisturbed cycles.

4.4. Determination of the Number of Clusters for *k*-Means Clustering

As shown in Figure 5, the scatter diagram shows the distribution of cycles start between 1:00 a.m. to 6:59 a.m. Notice that there is a cycle, represented within a red circle for illustration purposes that is far from other cycles in the diagram. The significant difference of this cycle indicates it as an outlier, while in order to maintain the integrity and continuity of the dataset, we did not remove this cycle from the original dataset. Based on the subsequent clustering result, the existence of this outlier affects the clustering quality evaluation process while it does not affect the partition result. More details of the outlier's impact on the clustering is mentioned in the next subsection.

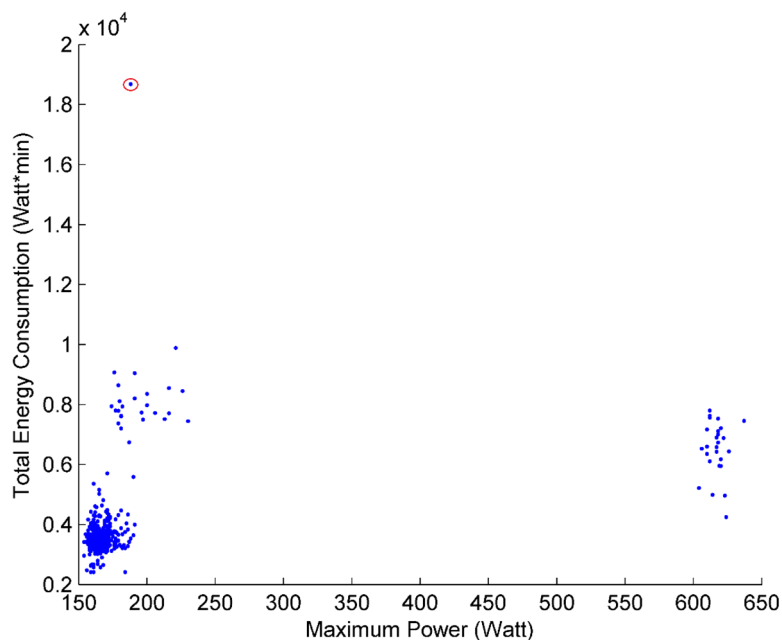


Figure 5. Distribution for cycles start between 1:00 to 6:59.

The distribution of the operation cycles indicates that they can be partitioned into several groups based on the selected attributes. To determine the precise number of clusters for the observed dataset, we compare the total error value of k that ranged from 1 to 10 based on the error function. Figure 6 plots the curve of total error with respect to k . Notice that the curve increases from point 1 to 2 and then has a significant drop at point 3. From point 4 to point 10, the curve different degrees of fluctuation, while none of them changed as significant as point 3. According to this trend, it can be understood that the first and most significant turning point is located at the point where k equals to 3. Meanwhile, the distribution’s shape depicted in Figure 5 indicates that the dataset could be clearly partitioned into 3 groups. Therefore, based on the total error curve for different k values and the distribution’s shape, the right number of clusters is specified as 3 for this study.

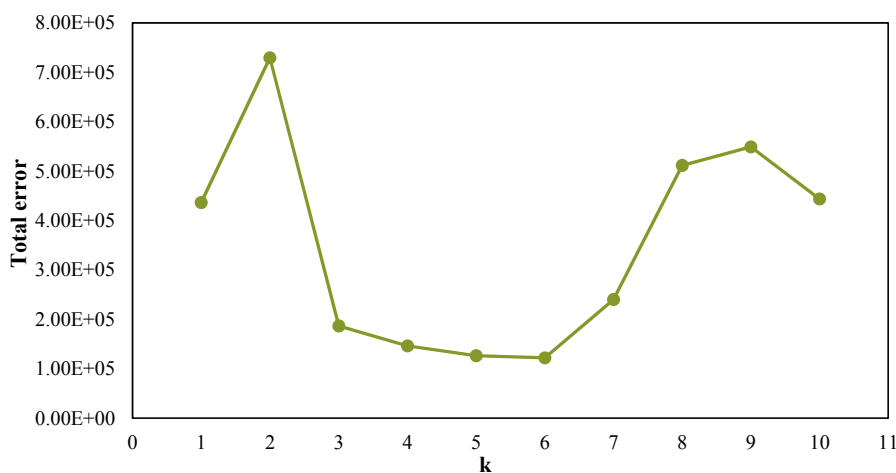


Figure 6. Total error curve.

4.5. Clustering Results

Once the number of clusters is determined, k -means clustering algorithm is applied to partition the dataset into the determined k clusters. Figure 7 depicts the k -means clustering result with k selected as 3. It can be seen from Figure 7 that different operation cycles are successfully partitioned as the boundaries of different clusters are clear.

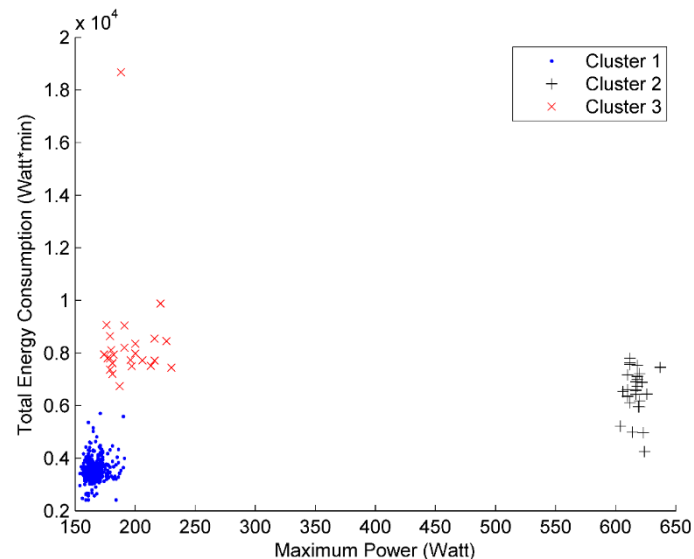


Figure 7. Clustering results for cycles start between 1:00 a.m. to 6:59 a.m.

An intrinsic method is then used to evaluate the clustering quality as the dataset is unlabeled and the ground truth is unavailable. Figure 8 depicts the bar chart of the silhouette coefficient value of operation cycles in cluster 1, 2, and 3. The silhouette coefficient bar chart is useful for measuring the clustering quality for each object. For operation cycles in cluster 1 and 2, the silhouette coefficient value is relatively stable and close to 1. This large value indicates that operation cycles in cluster 1 and 2 are close to cycles within the same cluster and far from cycles in other clusters. However, for cluster 3, the silhouette coefficient value for each operation cycle is unstable and relatively small as it contains an outlier.

To gain further insight into measuring clustering result quality, a statistical analysis of the silhouette coefficient value of all operation cycles in each cluster is conducted. Table 1 summarizes six statistical features of the silhouette coefficient value for each cluster. The six statistical features are the maximum value, the minimum value, mean, median, variance, and standard deviation. The average silhouette coefficient value over all operation cycles in a cluster indicates the compactness of the cluster. Notice that the mean silhouette coefficient value for cluster 1 and 2 are close to 1 and the variance is relatively small. This indicates a favorable clustering result as the cluster are compact and distinct. Because of the outlier, the performance of silhouette coefficient for cluster 3 is not as good as those for cluster 1 and 2. Since the outlier is far away from other operation cycles in cluster 3, the average distance of cycles within the same cluster has increased. Consequently, the average silhouette coefficient value for cluster 3 becomes small and the clustering performance for this cluster is less satisfied.

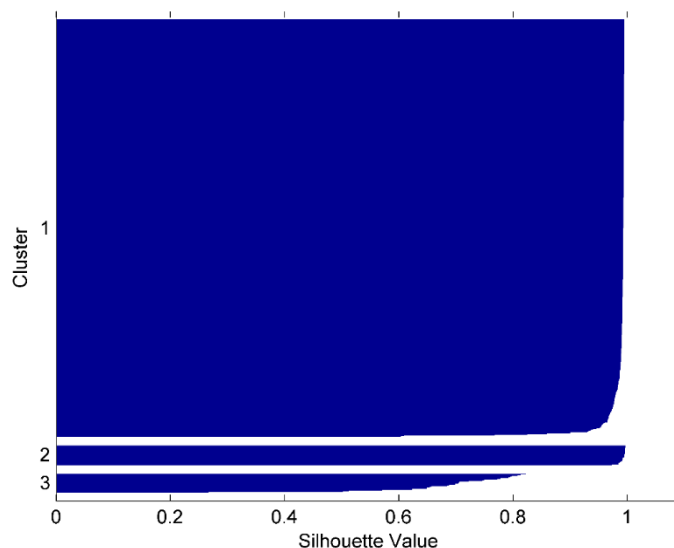


Figure 8. Silhouette coefficient value for different clusters.

Table 1. Statistical features of the silhouette coefficient values for each clusters.

Cluster No.	Max	Min	Mean	Median	Variance	Standard deviation
1	0.99419	0.60330	0.98682	0.99214	0.00076	0.02758
2	0.99589	0.97219	0.99191	0.99422	0.00003	0.00533
3	0.82238	0.26627	0.68672	0.70403	0.01491	0.12212

4.6. Identification of Classified Refrigerator Operation Cycles

Based on the clustering results, we summarize the characteristics of different clusters, such as total number of cycles, average maximum power, and average total energy use, in order to recognize and define different operation cycles. The results for different clusters are shown in Table 2. The start time for each cycle is also considered in this phase in order to find the sequential order between different clusters. Figure 9 illustrates the sequential order between different clusters for five days, from 18 January to 22 January. Notice that the relation between different days is disconnected as we only consider time period between 1:00 a.m. to 6:59 a.m. for each day. Based on the sequential order, it can be concluded that there is a high correlation between cluster 2 and cluster 3 as cycles in cluster 3 mostly occur right after cycles in cluster 2. Different clusters are defined as follows:

Table 2. Energy performance summary of different clusters.

Cluster No.	No. of cycles	Average maximum power (Watt)	Average total energy use (Watt * min)
1	576	165.92	3516.41
2	27	616.81	6558.74
3	26	194.12	8409.08

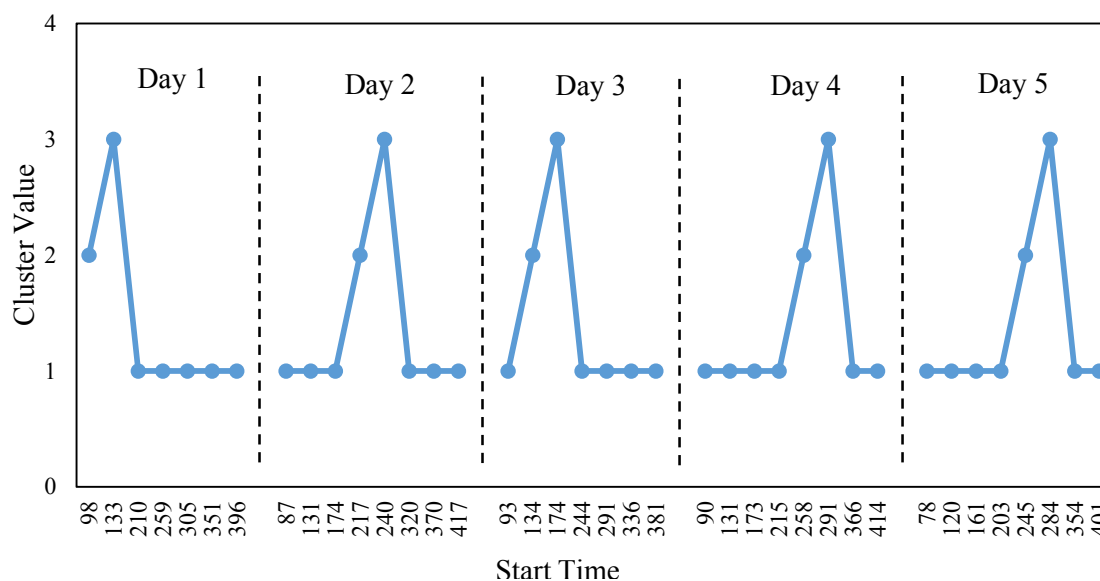


Figure 9. Cluster distribution from 18 January to 22 January.

Cluster 1: It can be seen from Table 2 that cycles in cluster 1 happen mostly (576 out of 629) and steadily. Their average maximum power which is the lowest among the three clusters indicates that the power of the refrigerator in cluster 1 is relatively small comparing with those in cluster 2 and cluster 3. As shown in Figure 9, the continuous appearance of cycles in cluster 1 implies that refrigerator is under stable operation condition and will maintain this condition until cycles in cluster 2 occur. Based on these features, cycles in cluster 1 are referred to the general working cycles which happen frequently and steadily.

Cluster 2: Cycles in cluster 2 are recognized as auto-defrost working cycles since they have the greatest average maximum power value which indicates the running of defrost heater. What's more, their frequency which is 27 out of 629 also demonstrates that refrigerator under these cycles is in an unordinary state.

Cluster 3: Cycles in cluster 3 are identified as auto-defrost affected working cycles as they occur right after auto-defrost cycles. In response to the auto-defrost cycles, these cycles use more time and consumes more energy than general working cycles to recover the refrigerator interior temperature.

5. Conclusions

An unsupervised machine learning model is proposed to classify household appliances operation cycles. The model was developed using k-means clustering as the classification algorithm and silhouette coefficient as performance measurement. To link each cluster to the specific operation mode, an identification process considering cluster characteristics related to operation mode is also provided. To validate the proposed method, a case study was conducted by testing it on a typical household refrigerator. In the case study, two attributes namely, maximum power and total energy use for each operation cycle were used for the classification. The undisturbed operation cycles are classified into three clusters and the silhouette coefficient performance indicated that each cluster is compact and distinct from others. The classification result shows that the proposed classification method could partition different types of operation cycle rationally by selecting proper attributes. The identification results confirm that the undisturbed refrigerator operation cycles can be recognized by considering the operation characteristics as well as the sequential order between different clusters.

The results of the case study demonstrates that the proposed classification method is capable of classifying operational cycles of household appliances, which can assist users in understanding the operation conditions of the appliances. The proposed method is beneficial for Smart Grid system as it provides a solution for the monitor to understand the way of operation of each appliances. In addition, the energy efficiency related index of each cluster, such as maximum power, total energy use, run time, and idle time can be used for refrigerator energy efficiency evaluation and fault detection. The identification of different operation cycles is significant in the research area of establishing appliance load profile, anticipating peak load, and intelligent control. Future work will focus on the research of pattern recognition for disturbed refrigerator operation cycles in order to comprehend the household refrigerator's energy performance and health, and subsequent integration with smart technologies.

Acknowledgments

The authors would like to acknowledge Jonathan Shi, Louisiana State University for providing data used in this study.

Author Contributions

Zeyu Wang designed the study, performed the data analysis, and wrote the manuscript. Ravi Srinivasan discussed results, reviewed and finalized the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Diamantoulakis, P.D.; Kapinas, V.M.; Karagiannidis, G.K. Big data analytics for dynamic energy management in smart grids. Big data research. *Comput. Sci.* **2015**, *2*, 94–101.
2. The SmartGrids European Technology Platform, The SmartGrids European Technology Platform. Available online: <http://www.smartgrids.eu/ETPSmartGrids> (accessed on 5 September 2015).
3. Yang, J.; Zhang, G.; Ma, K. A nonlinear control method for price-based demand response program in smart grid. *Int. J. Electr. Power Energy Syst.* **2016**, *74*, 322–328.
4. Cho, W.-T.; Chiu, Y.-S.; Wang, L.-C.; Lai, C.-F. A lightweight appliance recognition approach for smart grid. In Proceedings of the 2013 IEEE 11th International Conference on Dependable, Autonomic and Secure Computing (DASC), Chengdu, China, 21–22 December 2013.
5. Lai, Y.-X.; Lai, C.-F.; Huang, Y.-M.; Chao, H.-C. Multi-appliance recognition system with hybrid SVM/GMM classifier in ubiquitous smart home. *Inf. Sci.* **2013**, *230*, 39–55.
6. Srinivasan, D.; Ng, W.; Liew, A. Neural-network-based signature recognition for harmonic source identification. *IEEE Trans. Power Deliv.* **2006**, *21*, 398–405.
7. Chui, K.T.; Tsang, K.F.; Chung, S.H.; Yeung, L.F. Appliance signature identification solution using K-Means Clustering. In Proceedings of the IECON 2013—39th Annual Conference of the IEEE Industrial Electronics Society, Vienna, Austria, 10–13 November 2013.
8. Chu, Y.-H.; Qin, S.J.; Han, C. Fault detection and operation mode identification based on pattern classification with variable selection. *Ind. Eng. Chem. Res.* **2004**, *43*, 1701–1710.

9. Cheng, J.W.; Kendall, G.; Leung, J. Electric-load intelligence (E-LI): Concept and applications. In Proceedings of the 2006 IEEE Region 10 Conference TENCON, Hong Kong, China, 14–17 November 2006.
10. Huang, S.-J.; Hsieh, C.-T.; Kuo, L.-C.; Lin, C.-W.; Chang, C.-W.; Fang, S.-A. Classification of home appliance electricity consumption using power signature and harmonic features. In Proceedings of the IEEE Ninth International Conference on Power Electronics and Drive Systems, Singapore, Singapore, 5–8 December 2011.
11. Xu, W.; Dong, M. Tracking energy consumptions of home appliances using electrical signature data. In Proceedings of the 2012 IEEE Power & Energy Society General Meeting, San Diego, CA, USA, 22–26 July 2012.
12. Popescu, F.; Enache, F.; Vizitiu, I.-C.; Ciotirnae, P. Recurrence plot analysis for characterization of appliance load signature. In Proceedings of the 10th International Conference on Communications (COMM), Bucharest, Romania, 29–31 May 2014.
13. Kim, Y.; Kong, S.; Ko, R.; Joo, S.-K. Electrical event identification technique for monitoring home appliance load using load signatures. IEEE International Conference on Consumer Electronics, Las Vegas, NE, USA, 10–13 January 2014.
14. Chicco, G.; Napoli, R.; Piglione, F. Comparisons among clustering techniques for electricity customer classification. *Trans. Power Syst.* **2006**, *21*, 933–940.
15. Duan, P.; Xie, K.; Guo, T.; Huang, X. Short-term load forecasting for electric power systems using the PSO-SVR and FCM clustering techniques. *Energies* **2011**, *4*, 173–184.
16. Li, X.; Bowers, C.; Schnier, T. Classification of energy consumption in buildings with outlier detection. *IEEE Trans. Ind. Electron.* **2010**, *57*, 3639–3644.
17. Mastrullo, R.; Mauro, A.; Menna, L.; Palma, A.; Vanoli, G. Transient model of a vertical freezer with door openings and defrost effects. *Appl. Energy* **2014**, *121*, 38–50.
18. U.S. Energy Information Administration. *Estimated U.S. Residential Electricity Consumption by End Use*; U.S. Energy Information Administration, 2012. Available online: <http://www.eia.gov/tools/faqs/faq.cfm?id=96&t=3> (accessed on 12 November 2014).
19. Radermacher, R. Impact of environmental concerns on refrigerator/freezer designs. *Heat Recov. Syst. CHP* **1994**, *14*, 227–238.
20. Hasanuzzaman, M.; Saidur, R.; Masjuki, H. Investigation of energy consumption and energy savings of refrigerator-freezer during open and closed door condition. *J. Appl. Sci.* **2008**, *8*, 1822–1831.
21. Han, J.; Kamber, M.; Pei, J. Cluster analysis: Basic concepts and methods. In *Data Mining: Concepts and Techniques*, 3rd ed.; Morgan Kaufmann: Waltham, MA, USA, 2012.
22. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65.
23. PowerWise Systems. eMonitor—PowerWise System. Available online: <http://www.powerwisesystem-s.com/products/sitesage-emonitor4/> (accessed on 5 September 2014).