Classification and Analysis of User Behavior based on Action Duration in Residential Split-type Air Conditioners

Xiaobei Gu, and Meng Liu*

1School of Civil Engineering, Chongqing University, China

Abstract. In residential buildings, split-type air conditioners (ACs) are important indoor thermal environment control equipment. This study focused on ACs in China’s hot summer and cold winter zone. The valid data included 196 million real-time operation log records of more than 5,000 ACs in one year. The goal of this study is to classify action duration in a quantitative and interpretable way and analysis the user behavior based on the classification of action duration. Through the human intervention by indicators and multiple-Jenks algorithm, user behaviors based on action duration were divided into 6 categories and 12 types. The comparison of these action types in mode, wind speed and setting temperature were discussed as well. It reveals that some action durations are need to be focused in different conditions.

1 Introduction

In residential buildings, split-type air conditioners (ACs) are common equipment used for controlling indoor thermal environment, especially in China’s hot summer and cold winter (HSCW) zone, where the number of ACs per 100 households per year continues to increase [1]. Current studies related to the data-driven analysis of ACs behaviors has also been widely discussed [2]–[10].

In general, the data sources of ACs research are relatively traditional, such as survey interviews, experiments and on-site monitoring [2], [11], thus the sampling quantity and accuracy are limited to a certain extent. Moreover, the analysis methods and conclusions are relatively qualitative, focusing on local user behaviors, such as temperature settings or turning on-off [6], [7], [12]–[15]. Clustering methods are often used in behavior analysis. The common ones based on Euclidean distance, such as k-means, are suitable for multi-dimensional features. Feature extraction is not only a process of information loss, but also relatively subjective, because the direction of analysis results is related to clustering features [4], [6], [16], [17]. Specifically, detailed actions in ACs are rarely considered in clustering features, that is to say, the characteristic of behaviors is slightly lack of power.

The goal of this study is to classify action duration in a quantitative and interpretable way and analysis the user behavior based on the classification of action duration.

2 Methodology

2.1 Data source

The original data was the real-time tracking log uploaded by ACs in HSCW zone. About 196 million records from a whole year were extracted from the data warehouse, this study developed a Python module for data preprocessing. Finally, a total of 5099 effective devices were extracted, and the structured data was labelled as action layer, behavior layer, and user layer. Meanwhile, two working conditions of bedroom (wall-mounted) and living room (floor-standing) were considered separately.

2.2 Basic threshold of action duration

The basic threshold of action duration was selected by comparing the change trend of proportion of users ($P_n$) (1), usage probability ($AM_n$) (2) and weighted usage probability ($WM_n$) (3) in minutes, hours and days. The formulas are listed as follows:

$$P_n = \frac{c_n}{c_{users}}$$  

$$AM_n = \frac{\frac{2}{n} \sum_{i=1}^{n} y_i}{\frac{2}{n} \sum_{i=1}^{n} x_i}$$  

$$WM_n = \frac{\frac{2}{n} \sum_{i=1}^{n} x_i y_i}{\frac{2}{n} \sum_{i=1}^{n} y_i}$$

At the specific action duration $n$, $c_n$ - user amount, $x_n$ - action numbers. Others, $c_{users}$ - total user amount, $y_n$ - total action numbers.

* Corresponding author: liumeng2033@126.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
2.3 Multiple classification of action duration

2.3.1 Traditional Jenks natural breaks

Jenks natural breaks classification is a data clustering method which specifies an arbitrary grouping of the numeric data. The evaluation index is goodness of variance fit (GVF) (4) [18].

\[
GVF = 1 - \frac{SDCM}{SDAM}
\]

(4)

\(SDAM\) - sum of squared deviations from the array mean (constant), \(SDCM\) - sum of the squared deviations from the class means, the smaller \(SDCM\), the closer \(GVF\) is to 1, and the better index is, usually 0.8 would be considerable.

2.3.2 Multiple Jenks natural breaks

A more interpretable method was adopted since the Jenks result was invalid. Therefore, the GVF of improved Multiple-Jenks (5) was as follows:

\[
GVF_i = 1 - \frac{\sum SDBM_i}{kSDGM}
\]

(5)

\(SDSM\) - sum of the squared deviations from the set means, \(SDBM\) - sum of the squared deviations from the bin means, \(SDGM\) - sum of the squared deviations from the group means.

3 Results

3.1 Basic threshold of action duration

Comprehensive comparison of proportions of users, usage probability, and weighted usage probability as well as their logarithmic form in different time dimensions, the action duration was classified as 6 categories. The results are presented in Fig. 1, 2, and 3.

Fig. 1. The trend of three indicators and action duration in the dimension of minutes and hours between bedroom and living room.

Fig. 2. The trend of indicators and action duration in logarithmic in minute and hour dimensions in bedroom.

Fig. 3. The trend of indicators and action duration in logarithmic in minute and hour dimensions in living room.

3.2 Multiple classification of action duration

Based on the basic threshold, each category was divided with Jenks, and their original and multiple GVF was listed in Fig. 4 and 5.

Fig. 4. Independent classification by Jenks and comparison of four actions between bedroom and living room.
3.3 Validation of classification

With the validation in Fig. 6, the optimal solution of action duration is 10 (besides the QSA and BSA), the result is listed in Table 1.

| Name                                           | types       | Action Duration (WM) | Action Duration (FS) |
|------------------------------------------------|-------------|-----------------------|-----------------------|
| Quickly Setting the conditioner Action         | QSA-T1      | 000d00h00m - 000d00h01m | 000d00h00m - 000d00h01m |
| Belatedly Setting the conditioner Action       | BSA-T1      | 000d00h01m - 000d00h06m | 000d00h01m - 000d00h06m |
| Weakly Conditioning the environment Action     | WCA-T1/T2   | 000d00h06m - 000d02h15m | 000d00h06m - 000d02h15m |
| Strongly Conditioning the environment Action   | SCA-T1/T2/T3| 000d06h00m - 000d08h30m | 000d06h00m - 000d08h30m |
| Temporarily Maintaining the environment Action | TMA-T1/T2   | 000d16h00m - 001d04h30m | 000d16h00m - 001d04h30m |
| Permanently Maintaining the environment Action | PMA-T1/T2/T3| 002d00h00m - 028d00h00m | 002d00h00m - 021d00h00m |

4 Discussions

4.1 Mode comparison of actions

In general, mode comparison of actions is shown in Fig. 7, the cooling mode accounts for the largest proportions, and bedroom and living room are similar. The longer the action duration, the higher the proportion of auto mode and heating mode. Moreover, dry mode is basically proportional to auto mode.

4.2 Wind speed comparison of actions

In general, wind comparison of actions is shown in Fig. 8, the proportion of auto speed is far more than the quiet and low speed. In SCA (>6h), the quiet speed becomes larger than it in other type action, the possible reason is that the action occurs during sleep and no change at the moment.

4.3 Setting temperature comparison of actions

In cooling mode, setting temperature comparison of actions is shown in Fig. 9, the setting temperature of 26 accounts for the highest proportion. The shorter the action duration, the more uniform the proportion. In living room, the temperature distribution is higher with the increase of action duration.
5 Conclusions

In summary, the user behavior based on action duration in residential split-type ACs can be quantitively classified in to 6 categories and 12 types. Their proportions, mode, wind speed, setting temperature, and distribution were diverse. The result could be applied in the action selection in AC behavior analysis and dynamic thermal comfort experiment.

The research work was supported by the National Key R&D Program of China (Grant No: 2018YFD1100704), and the 111 Project of China (Grant No: B13041).

References

[1] China Statistics Press, National Bureau of Statistics of China, (2020).
[2] J. Du, W. Pan, and C. Yu, Energy Build., vol. 212, p. 109811, (2020).
[3] S. Hu, D. Yan, S. Guo, Y. Cui, and B. Dong, Energy Build., vol. 148, pp. 366–378, (2017).
[4] L. Yan, M. Liu, K. Xue, and Z. Zhang, J. Build. Eng., vol. 30, p. 101197, (2020).
[5] Z. Li, Heat. Vent. Air Cond., vol. 9, (2018).
[6] D. Xia, S. Lou, Y. Huang, Y. Zhao, D. H. W. Li, and X. Zhou, Energy Build., vol. 203, p. 109446, (2019).
[7] J. Yao, Energy Build., vol. 175, pp. 1–10, (2018).
[8] J. An, D. Yan, and T. Hong, Energy Build., vol. 174, pp. 214–227, Sep. (2018).
[9] L. Yan and M. Liu, Sustain. Cities Soc., vol. 60, no. June 2019, p. 102194, (2020).
[10] H. Liu, H. Sun, H. Mo, and J. Liu, Energy Build., vol. 250, p. 111297, Nov. (2021).
[11] D. Yan et al., Sustain. Cities Soc., vol. 46, p. 101393, (2019).
[12] X. Ren, D. Yan, and C. Wang, Build. Environ., vol. 81, pp. 172–182, (2014).
[13] H. Yoshino, T. Hong, and N. Nord, Energy Build., vol. 152, pp. 124–136, (2017).
[14] T. Ueno and A. Meier, Energy Build., vol. 228, no. February, p. 110423, (2020).
[15] X. Xu, W. Liu, and Z. Lian, Energy Build., vol. 220, p. 110086, (2020).
[16] A. Malik, N. Haghdadi, I. MacGill, and J. Ravishankar, Appl. Energy, vol. 235, no. FEB.1, pp. 776–785, (2019).
[17] M. K. Wright et al., Build. Environ., vol. 183, p. 107187, (2020).
[18] G. F. Jenks and F. C. Caspall, Ann. Assoc. Am. Geogr., vol. 61, no. 2, pp. 217–244, (1971).