Energy and Buildings 156 (2017) 121–133

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

Predicting hourly energy consumption in buildings using occupancy-related characteristics of end-user groups

Kwonsik Song [a], Nahyun Kwon [a], Kyle Anderson [b], Moonseo Park [a,*], Hyun-Soo Lee [a], SangHyun Lee [b]

[a] Department of Architecture and Architectural Engineering, Seoul National University, Room 39-425, Gwanak-ro 1, Gwanak-Gu, Seoul 151-742, Republic of Korea
[b] Department of Civil and Environmental Engineering, University of Michigan, 2350 Hayward St., 2340 G.G. Brown Building, Ann Arbor, MI, United States

ARTICLE INFO

Article history:
Received 31 May 2017
Received in revised form 25 August 2017
Accepted 20 September 2017
Available online 22 September 2017

Keywords:
Energy saving
Energy use prediction
Data mining techniques
Occupancy status
Occupants' energy use behavior

ABSTRACT

Accurate predictions of energy consumption are essential to optimizing building energy use performance. To date, substantial efforts have been undertaken to improve prediction accuracy, specifically while focusing on occupants' presence in buildings. Unfortunately, two significant obstacles remain when predicting building energy consumption using occupancy data. First, occupancy diversity among end-user groups is rarely considered during model development. Second, occupancy's correlation with energy consumption may be weak due to variances in occupant behavior. Therefore, this research aims to investigate how occupancy-related characteristics of end-user groups affect prediction performance. In order to achieve this objective, a data mining-based prediction model is constructed to mimic building thermal behaviors. The experimental results using the proposed prediction model make it evident that prediction accuracy is improved when considering diverse occupancy and its correlation with energy use. In addition, significant prediction accuracy is achieved using only a minimal amount of historical data. With the proposed prediction model, it is possible to obtain more detailed information about energy use patterns (e.g., load shape, the amount of energy use) for end-user groups. Thus, facility managers will be able to personalize the operation of energy-consuming equipment depending on end-user group for reducing energy consumption without compromising occupants' thermal comfort.

1. Introduction

As buildings consume 40% of all energy globally, improving their performance remains a critical task in order to meet energy saving goals [1]. Accordingly, much effort has been made to reduce the energy use in buildings. Being able to accurately predict energy use in buildings is essential to optimize the operation of energy-using equipment during a buildings operation [2–4]. Once it is understood where energy is consumed within a building, it is possible to develop appropriate energy saving strategies. This enables facility managers to achieve energy saving in the following three ways: 1) efficiently set starting/finishing time of heating, ventilating, and air conditioning (HVAC) systems, 2) avoid reaching peak energy demand by pre-heating/cooling of buildings, and 3) adjust heating and cooling setpoints during the peak energy periods.

In the extensive literature on building energy use prediction, various influential factors have been considered due to their significant correlation with energy consumption. These factors include: weather, building characteristics, equipment, and occupant-related characteristics [5–8]. Among these factors, recent studies have emphasized the importance of occupancy since occupants interact with energy-consuming equipment and devices in the built environment [8–12]. When attempting to predict building energy use, many previous efforts have employed fixed occupancy schedules as an alternative to using actual building occupancy data for simplicity and due to lack of readily available data [8,9]. However, smart monitoring systems now allow us to automatically obtain high-resolution occupancy data (e.g., 1 h time interval) and has begun to be used when constructing building energy use prediction models [10–13]. It has been found that considering occupancy as an input variable during model development improves prediction performance.

* Corresponding author.
E-mail addresses: woihj@snu.ac.kr (K. Song), prideknh@snu.ac.kr (N. Kwon), kyleand@umich.edu (K. Anderson), mspark@snu.ac.kr (M. Park), hyunslee@snu.ac.kr (H.-S. Lee), shdpm@umich.edu (S. Lee).

https://doi.org/10.1016/j.enbuild.2017.09.060
0378-7788/

Unfortunately, despite the recent advancements in prediction accuracy, two significant obstacles remain when predicting energy consumption using occupancy data. First, diverse occupancy, which refers to differences in occupancy status among end-user groups (EUGs), has rarely been considered during model development. This is important because many buildings have different EUGs which have different occupancy patterns. For instance, university buildings have complex occupancy patterns due to the varying functions of rooms: administration, research, lecture, and seminar [14]. Most studies [7–13] to date have eliminated occupancy diversity by averaging the values of occupancy status at the building level, which may contribute to discrepancies between actual and predicted energy use. Other studies have employed spatially granular data (e.g., floor and unit level) and individual equipment level data for building energy use prediction, but did not consider occupancy diversity during model development [2,15]. Second, occupancy's correlation with consumption may be weak at time due to variances in occupant behavior. If occupants fail to switch off their equipment and devices before leaving, energy will be consumed while unoccupied and uncorrelated with occupancy [16–19]. As a consequence, it can be difficult to ensure an improvement in prediction accuracy, i.e., the correlation effect. Most studies to date have used occupancy as an input variable without considering its correlation with energy use [11,12]. In rare studies that have investigated the correlation between energy use and occupancy status, the correlation effect remains unclear because there was no attempt to compare the performance of prediction models with different correlation levels [10,13].

Therefore, this research aims to investigate how occupancy-related characteristics of EUGs affect energy use prediction performance. In order to achieve this objective, a data mining-based prediction model is constructed since it facilitates to mimic building thermal behaviors [20] and identifies representative EUGs within buildings [21,22]. The developed model will provide more accurate information about daily peak demand and daily energy use in buildings. Furthermore, it is expected that the model makes it possible to recognize energy use patterns for EUGs. In turn, facility managers will be able to personalize the operation of energy-consuming equipment depending on EUG.

This paper is organized as follows. First, a literature review is presented on the application of occupancy data to building energy use prediction and prediction methodologies. Second, data mining techniques relevant to this work are introduced and discussed. Third, a data mining-based prediction model is developed by using real-world data collected from buildings in Seoul, South Korea. Next, the experimental results using the proposed prediction model are presented and the paper concludes with a discussion of the results followed by the conclusion.

2. Literature review

2.1. Application of occupancy data to building energy use prediction

In the extensive literature on building energy use prediction, occupancy data is typically substituted with building or equipment schedules which indirectly reflects the behavior of occupants. Kwok et al. [6] proposed a multi-layer perceptron model for building energy use prediction using the power consumption of primary air-handling units (PAU) as an alternative to occupancy data to mimic occupants' presence in a building. Yezioro et al. [8] constructed an artificial neural network (ANN) prediction model which uses the occupancy schedule as an input variable. In the optimized ANN model for building energy forecasting suggested by Li et al. [9], the opening schedules of a library were used to represent the hourly occupancy of each reading room.

Due to recent advancements in technology, it has become possible to monitor occupancy in real-time. With this new capability, researchers have begun using actual occupancy data to simulate occupants' behavioral characteristics in temporal and spatial contexts. Sandels et al. [10] presented a data analysis approach for conducting day-ahead predictions of electricity consumed by appliances, ventilation systems, and cooling equipment in an office building floor. It is found that the most significant predictor for the appliance load is the occupancy ratio. Virote and Neves-Silva [11] produced reliable predictions for building energy use by integrating stochastic occupant behavioral models with energy consumption models. Wang and Ding [12] proposed an occupancy-based energy consumption prediction model. The prediction model used a time-varying indoor occupancy rate which is obtained by using Monte Carlo simulation and Markov chain model. As a method to quantify energy savings by measurements and verification, Liang et al. [13] developed an energy baseline model using a short-time interval data on the number of occupants.

As mentioned above, a substantial number of studies investigated the effect of occupancy on the performance of building energy use prediction. However, the occupancy data used in previous studies was mostly simplified through aggregating occupancy status at the building level. Furthermore, when constructing prediction models, there were limited attempts to investigate occupancy's correlation with energy use. As a consequence of such significant obstacles, there still remains a discrepancy between the actual and predicted energy use.

2.2. Prediction approaches

Various building energy use prediction models have been proposed and can be categorized as: engineering, statistical, or ANN [23]. Engineering models simulate building energy use based on the physical and environmental factors [7,8]. While this approach has the advantage of being able to calculate elaborate thermal dynamics at a building level, it is not without limitation. This method can be a complicated difficult process which involves constructing a simulation model and obtaining meaningful input data [24]. Statistical models predict building energy use by using historical energy use data together with the measured input data [10,25,26]. While the structure of this approach is well understood due to the simplicity of the model parameters, substantial effort is required to overcome autocorrelation and multi-collinearity problems [4]. Lastly, ANN operates training procedures using historical data and then predicts building energy use [4–9]. This approach is highly applicable for solving non-linear and complex problems [27]. Not without limitations as well, ANN face potential problems with the reliability and accuracy of prediction results since it relies on the training data and can be computationally intensive [20,28].

As described thus far, each approach has its own advantages and disadvantages. Nevertheless, among these approaches, ANN prediction models are becoming increasingly more common in the field of building energy use prediction because the thermal behaviors of a building involve a non-linear problem [20].

3. Methodology

For this research, three data mining techniques are employed to predict the total amount of building energy use. First, *k*-means algorithm is used to investigate representative EUGs within buildings due to its ability to categorizes a set of objects into meaningful groups [29]. Second, artificial neural networks are constructed to predict energy use for the identified EUGs because it facilitates to