Energy and Buildings 84 (2014) 607–616

Contents lists available at ScienceDirect

Energy and Buildings

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

A new methodology for building energy performance benchmarking: An approach based on intelligent clustering algorithm

Xuefeng Gao*, Ali Malkawi

Department of Architecture, University of Pennsylvania, Philadelphia, PA, 19104, USA

ARTICLE INFO

Article history:
Received 4 June 2014
Received in revised form 18 August 2014
Accepted 19 August 2014
Available online 27 August 2014

Keywords:
Building performance
Energy benchmarking
Clustering algorithm
Multi-dimensional features

ABSTRACT

Though many building energy benchmarking programs have been developed during the past decades, they hold certain limitations. The major concern is that they may cause misleading benchmarking due to not fully considering the impacts of the multiple features of buildings on energy performance. The existing methods classify buildings according to only one of many features of buildings—the use type, which may result in a comparison between two buildings that are tremendously different in other features and not properly comparable as a result.

This paper aims to tackle this challenge by proposing a new methodology based on the clustering concept. The clustering concept, which reflects on machine learning algorithms, classifies buildings based on a multi-dimensional domain of building features, rather than the single dimension of use type. Buildings with the greatest similarity of features that influence energy performance are classified into the same cluster, and benchmarked according to the centroid reference of the cluster.

The proposed methodology contains four steps: feature selection, clustering algorithm adaptation, results validation, and interpretation. The experimentation was carried out with a comparison between the proposed methodology and the Energy Star approach. It was shown that the proposed methodology could account for the total building energy performance and was able to provide a more comprehensive approach to benchmarking. In addition, the multi-dimensional clustering concept enables energy benchmarking among different types of buildings, and inspires a new perspective to investigate building performance typology.

1. Introduction

1.1. Theory of building energy benchmarking

Originally, the word "benchmark" was used exclusively in topography to precisely define a reference point in terrain or geological analysis. In Merriam-Webster, benchmark means "a surveyor's mark. . .of previously determined position. . .and used as a reference point. . .standard by which something can be measured or judged." By this definition, benchmarking is a process used to measure something similar according to the previously determined reference point.

When the term was first used in building industry, it referred to energy benchmarking, specifically assessing the energy performance of buildings of similar type. Various studies followed different steps to conduct building energy benchmarking. A qualifying benchmarking process should contain at least three steps: collect a reasonably large database of building samples, obtain the energy performance information of the candidate buildings, and conduct comparison analysis [1–4].

First, it is necessary to possess a database with information on the energy performance of a significant number of building samples. The database can be developed by measuring data on site, conducting surveys, or using simulation modeling. The second step is to gather the performance information of the candidate buildings that are to be benchmarked. Third, a comparative analysis of the energy use of the candidate buildings against the samples held in the database is conducted to provide benchmarking results in terms of energy performance.

1.2. State-of-the-art methods

Many research projects have been conducted since the emergence of energy benchmarking in building industry. Broadly, the existing energy benchmarking methods can be categorized into four categories: points-based rating systems, hierarchal and

* Corresponding author. Tel.: +86 13590541200.
E-mail address: frankgaoxf@gmail.com (X. Gao).

http://dx.doi.org/10.1016/j.enbuild.2014.08.030
0378-7788/

end-use metrics, statistical (or regression-based) approaches, and simulation model-based approaches [5].

Points-based rating systems are best exemplified by the US Green Building Council's Leadership in Energy and Environmental Design (LEED) rating system. They do not allow comparisons against other buildings, rather, they provide standards and guidelines to measure how efficient and environmentally friendly a facility is. A LEED score is made up of credits assigned for satisfying different criteria including energy efficiency and other environmental factors. However, the scoring system for building energy efficiency can be misleading.

Hierarchal and end-use metrics refer to the generation of benchmarks that link energy usage to climate and functional requirements. The idea is to begin the analysis at the whole building level and gradually move down to the underlying system and components level to find the performance data. For instance, the highest level data can be gross area and annual utility bills, and the second highest level can be percentage of use types and heating, ventilation and air conditioning (HVAC) system, then the lowest level can be zone energy uses. This method is useful for accounting for more of the differences in features affecting energy use. However, the type of data required is usually not readily available. One strategy is to sub-meter the system and component loads, following the hierarchical process [5].

In Statistical approaches, statistics for a population of similar buildings are used to generate a benchmark against which the building EUI is compared. This method requires a large dataset to produce a reasonably sized sample of comparative buildings. Cal-Arch, for example, queries its database for similar buildings in California and provides histograms and statistics for the distribution of the query results [2,6]. EPA's Energy Star, on the other hand, accounts for more of the differences between buildings through the use of regression models and normalization methods that are used to generate a ranking score based on energy efficiency ratios [7].

Simulation model-based approaches calculate energy benchmarks based on an idealized model of building performance. The simulation engine, EnergyPlus, is one of many energy modeling tools used in these approaches. Models have many uses in benchmarking. They have the advantage to be tweaked to account for a wide range of factors that contribute to variation in energy use. They can also be used to generate targets and compare design alternatives. A disadvantage to many users is that they are actually simulation models, and benchmarks based on simulation modeling may not be well calibrated to the actual buildings stock data [5].

Several building benchmarking programs have been developed based on the above methods globally. The US EPA's Energy Star program is based on historical energy consumption data and easily obtained information of nationwide buildings [7]. Australian building greenhouse rating is also derived from actual amount of annual consumption of energy similar to EPA's Energy Star program [8]. Developed by National University of Singapore, Singapore E-energy uses statistics of collected energy use and occupancy data [9]. In Montreal, a Canada energy rating system combines the information from utility bills with on-site measurements and computer simulations [8]. In a Danish energy labeling system, the influence of owners is accounted for and the scoring is based on comparison of water consumption, energy use and CO_2 emissions to other similar buildings [10].

From the above programs, the best-known and most technically robust building energy benchmarking one in the US is the Energy Star program [2,4,5,7]. Energy Star program is based on a regression model, which includes building type, floor area, energy use and location inputs as well as occupancy-related factors such as number of occupants, operating hours and number of computers. Location is used to obtain weather data for use in the model. The Energy Star score (0–100) is an estimate of how many similar buildings nationwide have higher energy use intensities. For example, an Energy Star score of 75 signifies that the building EUI is better than 75% of similar buildings nationwide. This tool is the most valuable initial screening tool available for national building energy use analysis.

1.3. Current barriers and proposed approach

Though many building energy benchmarking programs are available, they hold certain limitations [11–13]. As discussed previously, points-based rating systems are not based on performance data, which may be arbitrary and misleading. Simulation model-based approaches may not be well calibrated to the candidate buildings, and the calibration process requires much detailed inputs and time commitment. Hierarchal and end-use metrics require sub-metered data, which makes it hardly applicable to a large number of buildings. While statistical approaches are efficient to be implemented on a large building dataset, existing programs such as Energy Star do not provide a continuously robust benchmarking model as expected [14].

A group of researchers examined several Northeastern schools using the Energy Star Portfolio Manager [15]. They found the results could be counterintuitive. An older school with less services and amenities could achieve a higher score than a modern school with more energy efficient technologies. Similar studies showed buildings with large data processing centers or buildings operating 24 h appeared to be very inefficient. These findings indicate other features rather than use type, such as equipment and operation schedule, also influence building energy performance.

More studies were conducted identifying other energy performance factors using regression models [8,16]. In a study, researchers calculated that a single story commercial building in an equatorial climate with $12 \times 60\,m^2$ on plan would have an annual energy consumption of $242\,kW\,h/m^2$ per year. But using the same floor plan for a 10-stories building, the energy consumption would be reduced to $188\,kW\,h/m^2$ per year. Another study conducted by the same group in a colder climate region in Turkey showed the impact of building form on heating energy use. These studies suggest that many other building features such as building height and form would also affect building energy performance. Therefore, a new methodology for building energy benchmarking should be studied, which should fully consider the impacts of multiple features on building energy performance.

To overcome the barriers in existing building energy benchmarking methods, it requires a way to classify and benchmark buildings based on the proximity measure of the multi-dimensional features, instead of the one-dimensional single feature 'use type'. It reflects on certain classification concept. In a book, the author pointed out that classification is a procedure of assigning a data item to a predefined set of categories [17]. Clustering shows the desirable relations of the dataset, and produces initial categories during the classification process. Clustering is considered one of the most useful approaches in data mining process since it can be used to discover groups and identify interesting distributions and patterns of the underlying datasets [18].

Clustering addresses the problem of finding a structure in a collection of unlabelled data. These unlabelled data could be buildings that have not been classified into any group [19]. The structure defines the similarity in between these 'unlabelled' buildings, based on which the similar buildings are put into the same group. Thus, clustering is the process of organizing objects into groups whose members are similar in some way, and a cluster is a collection of objects (buildings) which are "similar" between them and are "dissimilar" to the objects (buildings) belonging to other clusters