Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 110 (2017) 371 – 376

Energy Procedia

1st International Conference on Energy and Power, ICEP2016, 14-16 December 2016, RMIT University, Melbourne, Australia

Developing a hybrid model of prediction and classification algorithms for building energy consumption

Saeed Banihashemi[a]*, Grace Ding[a], Jack Wang[b]

[a]*School of the Built Environment, University of Technology Sydney (UTS), 15 Broadway, Ultimo 2007, Australia*
[b]*Faculty of Engineering and IT, University of Technology Sydney (UTS), 15 Broadway, Ultimo 2007, Australia*

Abstract

Artificial intelligence algorithms have been applied separately or integrally for prediction, classification or optimization of buildings energy consumption. However, there is a salient gap in the literature on the investigation of hybrid objective function development for energy optimization problems including qualitative and quantitative datasets in their constructs. To tackle with this challenge, this paper presents a hybrid objective function of machine learning algorithms in optimizing energy consumption of residential buildings through considering both continuous and discrete parameters of energy simultaneously. To do this, a comprehensive dataset including significant parameters of building envelop, building design layout and HVAC was established, Artificial Neural Network as a prediction and Decision Tree as a classification algorithm were employed via cross-training ensemble equation to create the hybrid function and the model was finally validated via the weighted average of the error decomposed for the performance. The developed model could effectively enhance the accuracy of the objective functions used in the building energy prediction and optimization problems. Furthermore, the results of this novel approach resolved the inclusion issue of both continuous and discrete parameters of energy in a unified objective function without threatening the integrity and consistency of the building energy datasets.

Peer-review under responsibility of the organizing committee of the 1st International Conference on Energy and Power.

Keywords: Energy optimisation; machine learning; prediction; classification; residential buildings.

1. Introduction

The potential to save energy by systematic building management and optimisation is known to be significant and

* Corresponding author. Tel.: +61-406-639380.
E-mail address: Seyed.S.BanihashemiNamini@student.uts.edu.au

Peer-review under responsibility of the organizing committee of the 1st International Conference on Energy and Power.
doi:10.1016/j.egypro.2017.03.155

it could be estimated from 5% to 50% [1]. With this respect, numerous machine learning methods have been applied in the recent decade for predicting, classifying and optimising the energy consumption of buildings focusing on the different important parameters but each optimisation algorithm requires specific objective function as its fundamental unit to model the parameters and minimise the values. In the energy optimisation context, developing the most appropriate objective function is a critical task as much as the reliable energy dataset generation particularly where both types of qualitative and quantitative data are involved with the optimisation solver. It can be inferred from literature that two types of machine learning algorithms; namely prediction [2] and classification [3] are run for continuous and discrete parameters of building energy consumption, respectively. However, there is a salient gap on the investigation of hybrid objective function development for energy optimization problems including qualitative and quantitative datasets in their constructs. Although in some cases, transformation techniques can be utilized for converting the continuous and discrete variables into each other [4] but some shortcomings such as losing integrity or randomness may be arisen [5]. Therefore, this study is to introduce a hybrid energy objective function covering both categorical and continuous data in the unified approach in which two well-known algorithms of Artificial Neural Network (ANN) and Decision Tree (DT) were developed integrally in order to find the best solution for energy optimisation functions and enhance the accuracy of data-driven energy modelling and prediction.

1.1. ANN & DT

ANN is a mathematical or computational model that tries to simulate the structure or functional aspects of biological neural networks. One of the applications of ANN in engineering field is to predict the outcome of non-linear statistical problems which is usually utilized to model complex relationships between inputs and outputs or to find patterns in datasets [6]. The thermal equations used to analyses and calculate energy loads are complex, making ANN a good platform to be used for this purpose. In this form, the network is presented with datasets obtained from simulations and the values of inputs are fed into each neuron or nod. The weights are then adjusted through learning algorithms iteratively until a suitable output is produced. A suitable output, in this case, suitable predicted annual energy load is the one which is as close as to the simulation results.

DT is also considered as the most applied type of machine learning algorithms in classification problems thanks to its wide use in practice [7]. The reputation of this algorithm is largely hinged to its interpretability and accuracy in delivering predictive models with understandable structure which generates useful information on the corresponding domain. In addition, DT is capable of processing both numerical and categorical parameters. However, this method is more appropriate and accurate in handling the categorical parameters rather than numerical data [3]. There are different types of DT algorithms including Simple Tree, Medium Tree, Complex Tree and Bagged Trees which follow the similar fundamental principles but different degrees of complexity in combining Trees. It applies a flowchart like tree structure to separate the dataset into different predetermined categories for presenting the interpretation, categorisation and generalisation on data [7]. With reference to the mentioned characteristics of ANN and DT, a comprehensive dataset should be collected for feeding the algorithms with inputs and outputs.

2. Dataset development

For data collection purposes, a four-story building consisting of four units on each floor was selected for simulation representing the conventional type of low-rise residential apartments. Each level area is 400 m^2 summing up to the total area of 1600 m^2. The building was modelled in Rhino 5, parameterised in Grasshopper software and simulated by EnergyPlus for annual energy estimation. For energy simulation, building calculation program was set to low-rise apartment and kitchen, bedroom, bathroom and dining room in each unit were defined as a zone which each zone had its own thermal properties. This approach enables the thermal engine to precisely quantify adjacencies and inter-zonal connections [8]. The thermostat was set between 18-26 ºC to provide thermal comfort for occupants and activate HVAC devices below or above this range. Four cities of Sydney, Moscow, Kuala Lumpur and Phoenix were chosen as representatives of Temperate, Cold, Tropical and Hot-arid Climates, respectively. The procedures taken for the data collection were resulted in generating 4435 of datasets including 13 inputs (the variables) leading to the output (annual energy consumption) [9] consisting of 1053, 1138, 1114 and 1130 data for the cities of Sydney, Phoenix, Kuala