Energy and Buildings 131 (2016) 132–141

Contents lists available at ScienceDirect

Energy and Buildings

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

Indoor occupancy estimation from carbon dioxide concentration

Chaoyang Jiang[a], Mustafa K. Masood[a], Yeng Chai Soh[a,*], Hua Li[b]

[a] School of Electrical and Electronic Engineering, Nanyang Technological University, 639798 Singapore, Singapore
[b] School of Mechanical and Aerospace Engineering, Nanyang Technological University, 639798 Singapore, Singapore

ARTICLE INFO

Article history:
Received 26 July 2016
Received in revised form 31 August 2016
Accepted 1 September 2016
Available online 14 September 2016

Keywords:
Occupancy estimation
Moving horizon CO_2 data
Feature scaled extreme learning machine
Scaled random weight matrix
Local smooth
Global smooth
x-Tolerance accuracy

ABSTRACT

This paper developed an indoor occupancy estimator with which we can estimate the number of real-time indoor occupants based on the carbon dioxide (CO_2) measurement. The estimator is actually a dynamic model of the occupancy level. To identify the dynamic model, we propose the Feature Scaled Extreme Learning Machine (FS-ELM) algorithm, which is a variation of the standard Extreme Learning Machine (ELM) but is shown to perform better for the occupancy estimation problem. The measured CO_2 concentration suffers from serious spikes. We find that pre-smoothing the CO_2 data can greatly improve the estimation accuracy. In real applications, however, we cannot obtain the real-time globally smoothed CO_2 data. We provide a way to use the locally smoothed CO_2 data instead, which is available in real-time. We introduce a new criterion, i.e. x-tolerance accuracy, to assess the occupancy estimator. The proposed occupancy estimator was tested in an office room with 24 cubicles and 11 open seats. The accuracy is up to 94%percent with a tolerance of four occupants.

1. Introduction

Indoor occupancy information is an important part of home and office automation. It can be used as an input for the control of indoor lighting systems [1,2] and Heat, Ventilation and Air-conditioning (HVAC) systems [3,4]. Studies have shown that around one-third of the energy consumed in buildings can be saved using occupancy-based control [5–7]. Therefore, indoor occupancy estimation has become a popular area of research in recent years.

With multi-camera video and pattern recognition techniques, the number of indoor occupants can be accurately estimated [8–10]. But these type of methods are intrusive and cannot be widely used due to the expensive hardware cost and privacy concerns. A number of terminal-based methods have also been proposed in which occupants are required to use a device, such as the keyboard and mouse [11], smartphones with WIFI [12] or Radio Frequency Identification (RFID) tags [13,14]. With these methods, data security can be an issue. Also, they cannot detect occupants not using the devices. Consequently, it is imperative to find a non-intrusive and non-terminal-based method for indoor occupancy estimation.

This paper provides a new indoor occupancy estimator based on real-time CO_2 measurement. CO_2 sensors are available for standard HVAC systems and hence no additional hardware cost is required. More importantly, CO_2 sensors are non-intrusive and non-terminal-based. Next, we briefly review the current non-intrusive and non-terminal-based occupancy estimators.

1.1. Related prior work

Many non-intrusive and non-terminal-based types of sensors have been used for indoor occupancy estimation, such as pyroelectric infrared (PIR) sensors [15], ultrasonic sensors [16], and microphones [17]. All these works can easily determine whether a room is occupied but cannot estimate the exact number of occupants [18].

The number of indoor occupants can be predicted using the Agent Based Model [19] and inhomogeneous Markov chain model [20,21] based on the statistical information of the historical occupancy data. These methods require no real-time sensor data but quite a large amount of training data.

Recently, estimating the number of indoor occupants from environmental parameters such as temperature, humidity, pressure and CO_2 concentration is becoming popular. In past works, system identification technique [22] and machine learning techniques [1,23–26] has been used to construct the relation between environmental parameters and the occupancy level. The number of occupants was considered to be depending on the current state of

* Corresponding author.
E-mail addresses: chaoyangjiang@hotmail.com (C. Jiang), must0006@e.ntu.edu.sg (M.K. Masood), eycsoh@ntu.edu.sg (Y.C. Soh), lihua@ntu.edu.sg (H. Li).

http://dx.doi.org/10.1016/j.enbuild.2016.09.002
0378-7788/

the environmental parameters. However, the estimated occupancy level is not very accurate even when the maximum number of the occupants is small (the accuracy is less than 75% for a maximum of 4 occupants in [23,24]).

Amongst all the environmental parameters, CO_2 concentration is the one that most correlates with the number of occupants. The indoor occupancy level can be solved from the mass balance equation [27–30]

$$m\frac{dc_t}{dt} = -Q(c_t - c_t^{\text{sup}}) + Po_t \tag{1}$$

where m is the mass of indoor air, Q is the airflow rate, P is the CO_2 generation rate per person, o_t, c_t, and c_t^{sup} are the number of occupants, CO_2 concentration of indoor air and supply air at the time instant t, respectively. This physical model assumes that the indoor CO_2 concentration is uniform. In real applications, this method suffers from three issues: (1) unpredictable opening of doors and windows, (2) uncertainties involved with the CO_2 concentration measurements, and (3) error of the calculated differential due to inaccuracies in the CO_2 measurement (the accuracy of the commonly used CO_2 sensors is around a few tens ppm). An approximated explicit solution in terms of the number of occupants obtained from the mass balance Eq. (1) has been used in [28,29]. If using the approximated solution, we can avoid calculating the differential of CO_2 concentration but many other parameters are required to be measured. For more detail about the approximated solution, see [29]. However, the first two issues are not easy to overcome when using this method.

To estimate the number of occupants, a recent work investigated the dynamic CO_2 concentration of a small seminar room under various scenarios in terms of different ventilation flow rates and all potential changes of occupancy level [31]. The occupancy changes can be estimated via the analysis of the dynamic CO_2 concentration. However, this work only considered the case that the number of occupants was increasing, and this method is impractical for the large rooms in which the number of occupants may have a large change in a short time interval.

To achieve better occupancy estimation, one approach is to combine CO_2 concentration with other types of parameters. In [3], a linear model of occupancy level in terms of CO_2 concentration and temperature was constructed. This method was tested in a room with a maximum of 4 occupants. The number of occupants was estimated with a 5 min time delay with an accuracy of up to 88.8%. In [32], the measurement of CO_2 was combined with indoor power consumption, PIR sensors and a microphone for occupancy estimation. The testing average error was 0.19 occupants for a room with maximum 4 occupants. To the best of our knowledge, most of the current work deals with very few occupants [3,23,24], and no non-intrusive and non-terminal-based method has been shown to be effective for rooms with more than 20 occupants.

1.2. Statements of our contributions

In this paper, we present an estimator which can tell us the number of real-time indoor occupants. The estimator is actually a dynamic model of the occupancy level in which the number of occupants at a particular time instant depends on the CO_2 concentration and the estimated occupancy level in a past time horizon. We have conducted an experiment in a 9.3 m × 20 m air-conditioned office room with 24 office cubicles and 11 open seats. The experimental results have verified the effectiveness of the proposed estimator.

To identify the occupancy estimator, we proposed a variation of the Extreme Learning Machine (ELM), called *feature scaled ELM* (FS-ELM). We added a feature layer for the standard ELM and set the feature-to-hidden layer weight matrix as a scaled random matrix. The FS-ELM retains the computational efficiency of the standard ELM and is shown to be more effective to identify the occupancy estimator. The application of the FS-ELM can be easily extended to other problems.

The measured CO_2 concentration sometimes suffers from serious spikes, which has a negative influence on the occupancy estimation. We found that smoothing the CO_2 data can greatly improve the performance of the occupancy estimator. To smooth the CO_2 data, the global information of CO_2 concentration in the time domain is required. However, when using the estimator, the real-time globally smoothed CO_2 data is unavailable because the future measurements are unknown. We used the locally smoothed CO_2 data to replace the globally smoothed data and provided one way to remove the accumulated error deduced by the replacement.

In addition, we introduce the notion of *x-tolerance accuracy* to assess the results of the occupancy estimator, which is a useful performance index when the number of occupants is large.

1.3. Organization of the paper

The rest of this paper is organized as follows. In Section 2, we introduce the problem of indoor occupancy estimation from CO_2 measurements. In Section 3, we present the FS-ELM with which the proposed occupancy estimator is identified. In Section 4, we identify the occupancy estimator from globally smoothed CO_2 data, and provide a way to use the estimator based on the real-time locally smoothed CO_2 data. We then show the experimental verification in Section 5, and the conclusions are given in Section 6.

2. Indoor occupancy estimation from CO_2 data

Various dynamic models of the indoor CO_2 concentration have been discussed in the literature [3,22,27–29,31]. In summary, we introduce the following generalized discrete-time state space model

$$c_k = g(\mathbf{c}_{k-l:k-1}, \mathbf{o}_{k-l:k}, \mathbf{v}_{k-l:k}) \tag{2}$$

where c_k is the CO_2 concentration around an indoor sensor node at time instant t_k, $g(\cdot)$ is a unknown function, and $\mathbf{c}_{k-l:k-1} = [c_{k-l}, c_{k-l+1}, \ldots, c_{k-1}]^{\mathrm{T}}$ is the sequence of CO_2 concentration at the past time horizon $[t_{k-l}, t_{k-1}]$. Similarly, $\mathbf{o}_{k-l:k} = [o_{k-l}, o_{k-l+1}, \ldots, o_k]^{\mathrm{T}}$ and $\mathbf{v}_{k-l:k} = [v_{k-l}, v_{k-l+1}, \ldots, v_k]^{\mathrm{T}}$ are the sequence of indoor occupants number and venting level, respectively.

In [3,22,27–29,31] the indoor CO_2 concentration is assumed to be uniform, and in [22,28,29,31] the length of time horizon $l = 1$, which implies that indoor CO_2 concentration is with Markov property (memoryless property). However, [33] showed that the gradient of indoor CO_2 concentration can be very large, and CO_2 emitted by certain occupants cannot be immediately sensed. Therefore, it is unreasonable to simply set $l = 1$ unless the sampling time is large enough.

Apparently, if the CO_2 dynamic model has been identified, the number of indoor occupants can be estimated based on the real-time CO_2 measurements. In [3], the estimator is designed by solving a deconvolution problem. In this strategy, the deconvolution process suffers from truncation errors and the error of the identified CO_2 dynamic model. Therefore, we intend to directly identify the occupancy estimator, thus avoiding the deconvolution process.

Considering the CO_2 dynamic model (2), one generalized occupancy estimator can be described as

$$o_k = f(\mathbf{c}_{k-l:k}, \mathbf{o}_{k-l:k-1}, \mathbf{v}_{k-l:k}) = f(\mathbf{x}_k) \tag{3}$$

where $f(\cdot)$ is the model to be identified, and

$$\mathbf{x}_k = [\mathbf{c}_{k-l:k}^{\mathrm{T}} \quad \mathbf{o}_{k-l:k-1}^{\mathrm{T}} \quad \mathbf{v}_{k-l:k}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^n \tag{4}$$