Energy & Buildings 174 (2018) 372–379

Contents lists available at ScienceDirect

Energy & Buildings

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

Building occupancy modeling using generative adversarial network

Zhenghua Chen, Chaoyang Jiang*

School of Electrical and Electronics Engineering, Nanyang Technological University, 50 Nanyang Ave, Singapore

A R T I C L E I N F O

Article history:
Received 7 April 2018
Revised 26 May 2018
Accepted 15 June 2018
Available online 4 July 2018

Keywords:
Building occupancy modeling
Generative adversarial network
Neural networks

A B S T R A C T

Due to the energy crisis and the awareness of sustainable development, the research on energy-efficient buildings has increasingly attracted attention. To achieve this objective, one important factor is to capture occupancy properties for building control systems, which refers to occupancy modeling in buildings. Due to the complexity of building occupancy, previous works try to simplify the modeling with some specific assumptions which may not always hold. In this paper, we propose a Generative Adversarial Network (GAN) framework for building occupancy modeling without any prior assumptions. The GAN approach contains two key components, i.e. a generative network and a discriminative network, which are designed as two powerful neural networks. Owing to the strong generalization capacity of neural networks and the adversarial mechanism in the GAN approach, it is able to accurately model building occupancy. We perform real experiments to verify the effectiveness of the proposed GAN approach and compare it with two state-of-the-art approaches for building occupancy modeling. To quantify the performance of all the models, we define five variables with two evaluation criteria. Results show that our proposed GAN approach can achieve a superior performance.

1. Introduction

Energy consumed in buildings accounts for around 40% of total energy usage in the world. Energy saving in buildings is an urgent task. Since occupants determine the cooling/heating load and ventilation requirement of an indoor environment, the number of occupants known as occupancy is vital for energy efficiency in buildings [1]. For example, the authors in [2] proposed an occupancy-driven control system which is able to save 15% of the total energy. Occupancy information also indicates the lighting requirement. If no occupants are present, all the lights can be automatically switched off to save energy. An adaptive lighting system was presented in [3] which demonstrates an energy saving of 35 – 75% for lighting systems in buildings. Moreover, the occupancy information can also be used in emergency egress for efficient evacuation and rescue [4]. Occupancy information is of great importance and with high complexity. In order to understand the properties of occupancy, in this paper, we mainly focus on occupancy modeling. The objective is to capture regular occupancy dynamics which can be leveraged to improve the performance of real-time occupancy estimation [5,13]. Moreover, occupancy models are able to generate occupancy time series for energy simulation tools, such as EnergyPlus [6] and DeST [7], which can simulate the energy performance of buildings before construction so that the facilities of buildings can be sized based on the simulation results. Since occupants are directly related to the energy consumption in buildings, occupancy dynamics is also widely used in enhancing the performance of energy forecasting [8].

Occupancy modeling is to model occupancy dynamics, meaning to understand the distribution of occupancy data. With this model, we can predict the number of occupants at any time instance without sensor observations. Many advanced occupancy models have been presented in prior works. Wang et al. proposed a probabilistic model for occupancy modeling in a single person office [9]. Two exponential distributions were leveraged to model occupied and vacant intervals. The simulation results showed that the vacant interval follows the exponential distribution, but the occupied interval violates it. Page et al. presented an inhomogeneous Markov chain model with two states, i.e. presence and absence, for occupancy modeling in single person offices [10]. They also defined a parameter of mobility to calculate the transition probability matrix of the inhomogeneous Markov chain model. In experiments, some important variables, such as the first arrival and last departure times of an occupant, were defined to evaluate the performance of their proposed model.

The occupancy in single person offices is relatively easy to be modelled. But the modeling of occupancy in multi-occupant rooms is more difficult. Richardson et al. presented a first-order

* Corresponding author.
E-mail addresses: chen0832@e.ntu.edu.sg (Z. Chen), chaoyangjiang@hotmail.com (C. Jiang).

https://doi.org/10.1016/j.enbuild.2018.06.029

inhomogeneous Markov chain model for occupancy modeling in a residential building [11]. They defined the state of the inhomogeneous Markov chain as the number of active occupants. This inhomogeneous Markov chain model is an extension of the model in [10] where the authors defined the state as the presence and absence of a single person office. Since the number of active occupants is small for residential buildings, the Markov transition probability matrix will not be very large, which makes their proposed approach applicable. But this case does not hold for commercial buildings where a relatively large number of occupants will share one room. For example, a room with 20 occupants will yield a transition probability matrix with a dimension of 21×21 (states from 0 to 20) which is difficult to solve. Hence, in [12], Chen et al. proposed a novel inhomogeneous Markov chain model for occupancy modeling in commercial buildings. In a multi-occupant room, they defined the state of the inhomogeneous Markov chain model as the increment of occupancy with the assumption that only one occupant moves into or out of a room in a short time interval. In this way, the transition probability matrix will be with a dimension of 3×3 regardless of the total number of occupants in the room.

Gunathila et al. proposed a generalized event-driven approach for building occupancy modeling [14]. They assigned each occupant into groups based on some similar properties. Then, the occupant behaviour is fully driven by group and personal events. In [15], Wang et al. also presented an event-driven approach for the modeling of building occupancy [15]. They divided one day into some events, i.e. walking around, going to the office, getting off work and lunch break. Within each event, a Markov chain model was utilized for simulating occupancy patterns. There are two limitations for these two event-driven approaches. The first one is that the definition of events is ambiguous, and some personal events, such as sick and emotion changes, are difficult to define. The second one is that these two works only apply simulation data for model verification instead of real occupancy data.

Another relevant work can be found in [16]. The authors presented an agent-based model with some intuitively defined modules for the modeling of building occupancy. For instance, they defined a damping process for occupants who are in their working places and an acceleration process for occupants who are in hallways or restrooms. These intuitively defined modules lack a theoretical support. Thus, their proposed approach has limited performance. Jia et al. proposed a queueing approach to model building occupancy [17]. The designed queues have an infinite number of servers with time-varying parameters. To deal with the problem of abrupt changes of occupancy, they developed a piecewise homogeneous queue with the adaptation of the length of each homogeneous piece based on occupancy variations. However, their proposed method is built upon the assumption that the occupied time follows an exponential distribution which violates the conclusion in [9]. A comprehensive review of the approaches for building occupancy modeling can be found in [18].

Due to the complex behaviour of occupants, the occupancy data is very difficult to model. Existing works for occupancy modeling based on Markov chains or agent-based schemes all require the data to follow some assumptions such as first-order Markov property. Recently, a new powerful generative model named Generative Adversarial Network (GAN) has been developed [19]. It has been successfully applied in many challenging research areas, such as image processing [20] and natural language processing [21]. Owing to the powerful modeling ability, the GAN is able to model complex data even with some implicit distributions, and it does not require the data to follow any specific assumptions. Therefore, it is naturally suitable for the task of occupancy modeling. In this paper, we propose a GAN framework for building occupancy modeling. Owing to the universal approximation ability of neural networks, the GAN approach is able to learn the implicit distribution of real occupancy data for occupancy modeling. The GAN approach consists of two components: a generative network and a discriminative network. Firstly, the generative network is able to produce an occupancy time series with some random inputs. Then, the discriminative network attempts to distinguish these generated occupancy time series from the real ones. The final objective is to learn these two networks so that the generative network can produce the occupancy time series with the same distribution of the real occupancy time series and the discriminative network is unable to separate the occupancy time series generated by the generative network and the real occupancy time series. We perform real experiments to evaluate the performance of the proposed approach for occupancy modeling. Moreover, a comparison with some state-of-the-art approaches has also been made using real experimental data.

The main contributions of this work are summarized as follows:

- We propose a GAN framework with two powerful neural networks for the complex task of occupancy modeling without any prior assumptions.
- We apply real experimental data to verify the effectiveness of the proposed approach.
- The proposed GAN framework outperforms the state-of-the-art techniques for occupancy modeling.
- We evaluate the impact of one key parameter of neural networks in the proposed GAN framework, i.e., the number of hidden nodes, on the performance of occupancy modeling.

The remaining of the paper is organized as follows: In Section 2, we introduce the basic theory of GAN, followed by the proposed GAN framework for building occupancy modeling. In Section 3, firstly, we present the data collection process and define some variables and criteria to quantify the performance of building occupancy modeling. Then, the experimental setup is presented. After that, we demonstrate the experimental results with some discussions. Finally, we investigate the impact of one key hyperparameter of the proposed GAN approach on modeling performance. In Section 4, we conclude this work and present some potential future works.

2. Methodology

In this section, we show how the GAN is used for occupancy modeling. Firstly, we briefly introduce the basic theory of GAN. Then we propose a GAN framework for building occupancy modeling.

2.1. Generative adversarial network

The generative adversarial network was first proposed by Goodfellow et al. in [19]. It intends to learn the distribution of the given data and then generates the data under the same distribution. More specifically, it contains two parts, i.e. a generative (G) network and a discriminative (D) network. For the G network, the input can be a noise vector $\mathbf{z}$ if no prior knowledge is available. The G network attempts to map the vector $\mathbf{z}$ to data space as $G(\mathbf{z}, \theta_G)$ where θ_G consists of the parameters for the G network, and $G(\cdot)$ is a type of neural network, such as a multilayer perceptron. For the D network, it can be treated as a classifier for a binary classification. Specifically, it tries to distinguish the data generated by the G network $G(\mathbf{z}, \theta_G)$ from the real data $\mathbf{r}$. We define $D(\mathbf{x}, \theta_D)$, where θ_D consists of the parameters for the D network, as the probability that the data $\mathbf{x}$ is from the real data instead of the one generated by the G network. Here, the function $D(\cdot)$ is a classifier. The main objectives of the GAN are: 1) to maximize the classification accuracy for the D network which means to correctly identify the