BUILD SIMUL (2011) 4: 149 – 167
DOI 10.1007/s12273-011-0044-5

Research Article

A novel approach for building occupancy simulation

Chuang Wang, Da Yan (✉), Yi Jiang

Department of Building Science, School of Architecture, Tsinghua University, Beijing 100084, China

Abstract

Building occupancy is an important basic factor in building energy simulation but it is hard to represent due to its temporal and spatial stochastic nature. This paper presents a novel approach for building occupancy simulation based on the Markov chain. In this study, occupancy is handled as the straightforward result of occupant movement processes which occur among the spaces inside and outside a building. By using the Markov chain method to simulate this stochastic movement process, the model can generate the location for each occupant and the zone-level occupancy for the whole building. There is no explicit or implicit constraint to the number of occupants and the number of zones in the model while maintaining a simple and clear set of input parameters. From the case study of an office building, it can be seen that the model can produce realistic occupancy variations in the office building for a typical workday with key statistical properties of occupancy such as the time of morning arrival and night departure, lunch time, periods of intermediate walking-around, etc. Due to simplicity, accuracy and unrestraint, this model is sufficient and practical to simulate occupancy for building energy simulations and stochastic analysis of building heating, ventilation, and air conditioning (HVAC) systems.

Keywords

building occupancy,
occupant movement,
stochastic process,
Markov chain,
energy simulation

Article History

Received: 27 April 2011
Revised: 26 May 2011
Accepted: 28 May 2011

Building Thermal, Lighting, and Acoustics Modeling

1 Introduction

Building energy simulation tools such as EnergyPlus (Crawley et al. 2001), ESP-r (ESRU 1999), DeST (Yan et al. 2008; Zhang et al. 2008), and TRNSYS (Klein et al. 2004) have been playing more and more important role in building energy conservation since their conception. In past decades, modeling of heat and mass transfer processes, ambient weather data, as well as HVAC (heating, ventilation, and air conditioning) systems in buildings have been the main focus and have been well established. They support the application of simulation technique to predict and evaluate the performance of buildings (indoor climate, energy consumption, etc.). Occupant behavior, as a basic factor in building performance, still remains a big issue because of its stochastic nature in time and space.

In general, the behavior of building occupants can be broken down into two aspects: one is how they occupy the building (when they occupy the building and how many people for each zone), which is denoted by *occupancy*. The other is how they interact with building devices including windows, doors, blinds, air conditioning terminals, lights, and equipment (TVs, computers, printers, etc.). In most situations, occupants have the right and means to adjust and control these devices, thus, those interactions are closely related to occupancy. For example, lights in a zone might be turned on by occupants when the zone is occupied and might be turned off if there is no occupant. In some smart buildings, occupancy sensors are installed for automatic control of devices according to the occupied status of the monitored space to reduce unnecessary energy use while maintaining the comfort level of the environment. For buildings with occupancy controls, occupancy becomes a key driving factor to accurately predict the energy consumption of the buildings or the impact of such occupancy-based control system on building energy performance.

The load calculation and performance analysis of building HVAC systems depend on accurate accounting of internal heat gains due to occupancy (and associated use of lights and equipment), especially for interior zones in large public

E-mail: yanda@tsinghua.edu.cn

or commercial buildings where the internal heat gains are the most important factor that affects the indoor thermal environment. In such situations, the uncertainty or stochastic variation of internal heat gains usually results in overcooling or overheating during HVAC system operation, and thus plays a decisive role in the performance evaluation of a HVAC system, such as the comparison between CAV (constant air volume) and VAV (variable air volume) system (Yan 2005).

Therefore, how the stochastic characteristic of occupancy in building can be defined and modeled is an important issue that remains to be addressed in building energy simulation.

Currently the most commonly used method in simulation tools to represent occupancy is the so-called *schedule, diversity profile* or *diversity factors* (DeST 2008; EnergyPlus 2009; Davis and Nutter 2010), which can describe the time-variation of occupancy. A daily schedule consists of 24 hourly values while a yearly schedule usually consists of 365 daily schedules, in which the hourly values can be estimated from individual experience or onsite survey; the same schedule is usually used for zones that have similar functions. Through this deterministic approach, the average impact of internal heat gains (from occupant, lighting and equipment) on energy consumption and cooling load of the building can be estimated, but it cannot represent the stochastic variations of occupancy in time and space. In addition, the datasets corresponding to occupancy schedule in all its stochastic variety from survey and measurement are still scarce, and even if such datasets were available, since occupancy may vary widely from one building to another depending on the type and size of the building, etc., they may not accurately represent actual occupancy.

Consequently, stochastic models have been proposed to produce synthetic occupancy schedules in a certain way as inputs to simulation tools. A fundamental stochastic model uses the Monte Carlo method (Macdonald and Strachan 2001), which generates an estimation of occupancy for each zone in a building based on the probability distribution of occupancy in each zone (normal distribution used in practice). It is an easy way to consider occupancy uncertainty, regardless of the interactions and relationships of the occupancy at different times in a zone and that in different zones at the same time in a building, which are distinct in reality (e.g., an occupant is likely to stay in the same zone at the next time if he stays there at previous time and he cannot appear at two places at the same time). Thus, this method is too rough for occupancy modeling and usually only used for risk analysis.

In recent years, several advanced models have been proposed to randomly generate plausible sequences of occupancy in buildings. Wang et al. (2005) proposed a probabilistic model to predict and simulate occupancy in single person offices, where non-homogeneous Poisson process model with two different exponential distributions are used to generate the occupancy sequence in a single person office. Based on the examined statistical properties of occupancy, the durations of presence and absence during business hours are both exponentially distributed and the coefficient of each exponential distribution for a single office can be treated as a constant over the workday. Meanwhile, in order to combine the clock-time information into the simulation, the morning arrival time, the night departure time, and the lunch break time are assumed to be normally distributed, which are actually not supported by observations. This method is simple and elegant for single offices, but it is severely limited in describing the relationship of occupancy in different zones due to occupant movement in a building, and thus very hard to extend into other situations.

Page et al. (2008) presented an approach based on the inhomogeneous two-state Markov chain, where the model can generate a time series of the state of presence (in/out, or, present/absent) of occupants within a specific zone of a building, and the transition probabilities of the model, corresponding to arriving, leaving and staying in the respective states, are time-dependent and estimated from the probability of presence (based on aggregate occupancy records) at every time step. In addition, an assumed parameter of mobility describes the probability of state changes. This model is capable of reproducing the important characteristic of occupancy in office buildings such as the morning arrival time, the night departure time, typical long absences and the effective time of presence of the occupant within the zone by a uniform Markov chain. However, it has two major disadvantages: (1) as inputs to the model, the profiles of probability of presence and parameters of mobility are too complex to specify in simulation and to obtain from survey or measurements because of their time dependency; (2) this model does not simulate the movement of occupants from one zone to another, which is distinct in reality and important for occupancy prediction. Even if such efforts can be made, extending the model to multiple zones is much more challenging for determining the time varying entries of high-order transition probability matrix and parameters of mobility for state changes.

The most recent model is proposed by Liao et al. (2011), an agent-based model of occupancy dynamics in a building. This model regards each occupant as an agent and decides the state of an agent (the location of each occupant) at every time step through a set of rules specified by four modules, which can then be collected to generate time-series of zone-level occupancy. Compared to Page's model, Liao's model maintains a Markov-like property of agent dynamics and is easily scalable to an arbitrary number of zones and an arbitrary number of occupants. The scheduled activities of agents, the zones that each agent can access, the maximum occupancy limits of zones, and a secondary agent that occupies the building for brief periods of time (like visitors) are also involved. This model's prediction accuracy was found to be quite good in the single-occupant single-zone scenario and multi-occupant single-zone scenario, but is poor in the multi-occupant multi-zone scenario. The most time-consuming part in constructing the model is specifying the nominal presence probability profile for each agent in the preliminary state generator module, which consists of the probability time series that an agent occupies a zone at every time step. Due to excessive information, these inputs for the model are hard to obtain especially for multi-occupant multi-zone scenarios. For this reason, although it has no limitation in theory when applied to multiple zones, some simplifications have to be made in practice when specifying the nominal presence profiles, such as the zones that each occupant can access only include a primary zone, a secondary zone, a hallway, and a restroom. The frequency and average duration of visiting a restroom per day are assumed to be constant over the workday. Lunch and dinner breaks are strictly specified by fixed schedule rather than random process. These simplifications reduce the capacity of the model and would lead to poor results of the model in a multi-occupant multi-zone scenario.

Additional models have tried to generate the stochastic occupancy in buildings based on occupant activity simulation (Tabak 2008; Goldstein et al. 2010), where the models focused on the chronological sequence of detailed activities, with attributes of a task (e.g., working, meeting, eating), such as frequency, duration, priority, location, facilities, interactions with one another, etc. This method is quite sophisticated and can output many details of human behavior in a building, from which the occupancy is a natural consequence. However, relying on a huge extensive survey or measurement of such detailed activity, this indirect method is not as practical for occupancy simulation as the aforementioned methods.

The above overview indicates that existing models have demonstrated the capacity to realistically reproduce key properties of occupant presence for both the single-occupant single-zone scenario and the multi-occupant single-zone scenario, but have serious limitations when extended to multiple spaces. Furthermore, the excessive inputs of these models due to time dependency or high order of matrix are complex to handle in simulation.

The principal goal of this paper is to present a novel approach for occupancy simulation based on homogeneous Markov chain and demonstrate it in an office building simulation. In this study, occupancy is handled as the straightforward result of occupant movement processes which occur among the spaces inside and outside the building. Thus there is no explicit or implicit constraint to the number of occupants and the number of zones. By using the Markov chain method to simulate this stochastic movement process, the model can generate the location for each occupant and the occupancy for each zone of the building while maintaining a simple, clear set of input parameters. From the case study, it can be seen that the model can produce the realistic occupancy variations in the office building for typical workday with key statistical properties of occupancy such as the time of morning arrival and night departure, lunch time, periods of intermediate walking-around, etc. In the terms of simplicity, accuracy and unrestraint, this model is sufficient and practical to simulate occupancy for building energy simulations and stochastic analysis of building HVAC systems.

2 Methodology

The goal of the model in this paper is to generate stochastic occupancy schedules with the same statistical characteristic of building occupancy, which can then be used by a building energy simulation tool to further estimate the energy consumption of a building and the performance of a HVAC system.

The basic idea of the model is that building occupancy is a straightforward result of occupant movement processes which occur among the spaces inside and outside the building. The first-order homogeneous Markov chain technique, a widely-used well-established stochastic process method (Ross 1996), was selected for simulating the occupant movement process.

The concept of the first-order homogeneous Markov chain (HMC) technique is that any future state is dependent only on the present state together with the probabilities of the state changing (called Markovian property). These probabilities are held in the transition probability matrix and are time-independent (i.e., fixed). Consider a stochastic process $\{X_k, k=0,1,2,\cdots\}$ that takes a set of nonnegative integers $I=\{0,1,2,\cdots\}$ as possible values, HMC (with discrete states and discrete time steps) can be present by

$$\begin{aligned}&\boldsymbol{P}\{X_{k+1}=j \mid X_k=i, X_{k-1}=i_{k-1},\cdots,X_1=i_1,X_0=i_0\}\\&=\boldsymbol{P}\{X_{k+1}=j \mid X_k=i\}=p_{ij}(k)=p_{ij}\end{aligned}$$

for all states $i_0,i_1,\cdots,i_{k-1},i,j$ and all time steps $k \geqslant 0$. The fixed value p_{ij} represents the probability that the process will, when in state i, next make a transition into state j. For the transition probability matrix (denoted by $\boldsymbol{P}$ matrix in the sequel) that consists of one-step transition probabilities p_{ij}, we have that

$$p_{ij} \geqslant 0,\ i, j \in I;\ \ \sum_{j \in I} p_{ij} = 1,\ i \in I \tag{1}$$

HMC has many good statistical analysis features which greatly help set up the model. The interested reader can refer to Ross (1996) for more details on Markov chain.

2.1 General description

The proposed model has a two-level hierarchical structure consisting of a basic module named *movement process* and a high-level module named *events* as shown in Fig. 1. The module of movement process essentially implements a simulation of the Markov chain process and generates the locations of occupants step by step, which can then be used to calculate the occupancy for each zone in buildings. The module of events is used to specify the transition probabilities of Markov chain in specific periods of time, in order to represent the occurrences associated with time.

2.1.1 Movement process

In this paper, the process of occupant movement covers all the occurrences that correspond to the location change of people within a building, such as entering or leaving a specific space, moving around from one space to another, going outdoors for a while, etc. Such a stochastic process results in the variation of building occupancy in time and space. Such processes also may vary widely from one building to another, depending on the type and size of the building, the geographic location and climate, the ethnicity and preferences of the occupants, etc.

Consider a building with n zones that is occupied by m individuals, where a *zone* is an internal space in the building and indexed as $1,2,\cdots,n$; an individual is denoted by *occupant*, i.e., residents in case of a residential building or office workers in case of an office building, etc. The outside of the building is also involved and treated as a specific space indexed by 0, to form a complete movement graph.

Regarding the location of an occupant (in which space the occupant is) at every time step as a random variable, its possible values belong to the set (or a subset) of all spaces' indices $\{0 = \text{outside},\ 1 = \text{zone } 1,\ 2 = \text{zone } 2, \cdots, n = \text{zone } n\}$ that correspond to the occupant's accessible range. The movement process of each occupant can thus be described by a Markov chain in which the state of the process is exactly the location of an occupant, and the next location of the occupant is dependent only on the present location and the fixed transition probabilities held in the $\boldsymbol{P}$ matrix.

Here the following assumptions have been made: (i) the location of occupant due to movement has a Markovian property; (ii) any location change of occupant due to movement can be finished in one time step; (iii) the movements of each occupant are independent, thus each occupant has his own transition probability matrix.

The assumption (i) is supported by the experiments of Wang et al. (2005) and Dodier et al. (2006). They measured the occupancy states for some single offices by assembling a sensor network. If the location of an occupant is described by a HMC, the sojourn time of the occupant in any state (i.e., presence/in his office, absence/not in his office during working hours) should be geometrically distributed. From their results, it is approximately geometric for both presence and absence durations of an occupant in business hours. Figures 2 and 3 show the fitted and observed probability distribution of durations of presence and absence for an office in Wang's experiment, with a time interval of 15 minutes. Figures 4 and 5 show the histograms of sojourn times and the fitted probability distribution of durations of presence and absence for two offices in Dodier's experiment, with time intervals of 500 s and 200 s. Both authors concluded that the presence and absence durations are exponentially distributed. However, since the observed data in the experiments are analyzed with a discrete time interval, the exponential distributions are indeed geometric distributions, in a discrete form.

The two experiments, although based on single offices, in a way validate the assumption (i) in our model. It further ensures that the stochastic occupancy model based on occupant movement is reliable in other situations, even for multi-occupant multi-zone scenarios.

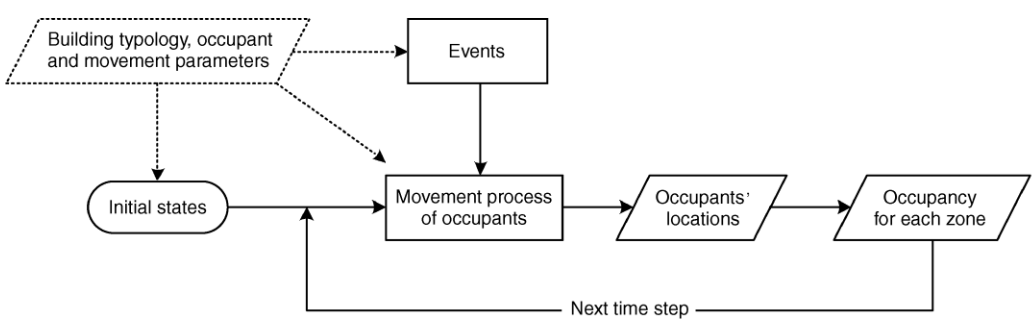


Fig. 1 The schematic of the model

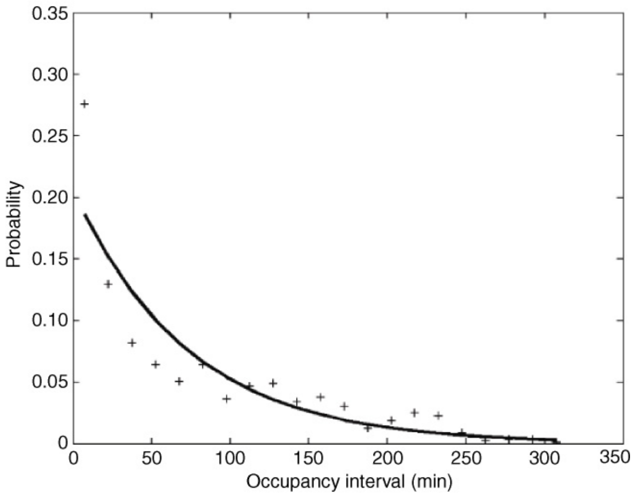


Fig. 2 Fitted and observed probability distribution of the presence durations (Wang et al. 2005)

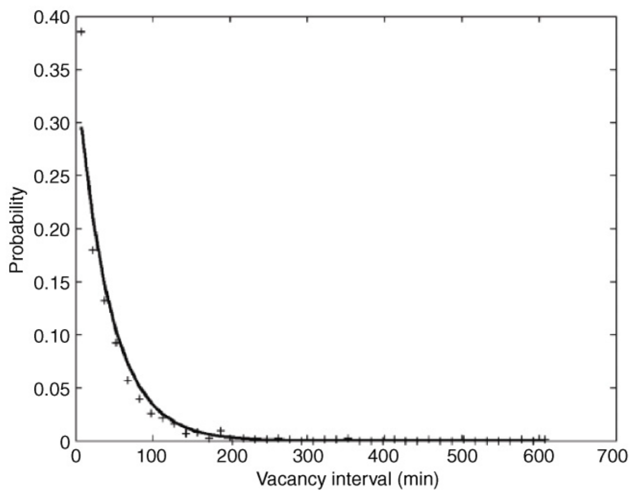


Fig. 3 Fitted and observed probability distribution of the absence durations (Wang et al. 2005)

As for the hypothesis (ii), since the movement of an occupant is essentially a continuous-time process, whether the location change of an occupant governed by ***P*** matrix can be realized depends on his movement speed, especially in a small time step simulation. The transition probabilities of the occupant from the present state to the next state are actually influenced by the limit of movement speed. For the moment, it is not taken into account. Accordingly, to ensure the temporal resolution of the results and to easily integrate with building energy simulation tools, the time intervals for the model can be 5, 10, 12, 15 minutes, or any other submultiples of 60 minutes (i.e., 1 hour that is the usual time step for building energy simulation).

With the above hypotheses, the movement patterns of each occupant can be modeled individually in a simple way even if they may be quite different, and the realistic Markovian properties of occupant movement can be reproduced.

2.1.2 Event mechanism

It is noted that such a HMC can only produce a pure random movement process for each occupant among the spaces (inside and outside of the building), without considering the time factors for the movement occurrences that probably happen for occupants in certain periods of time. Take an office building as an example, in normal conditions the employees would usually go to the office in the morning and leave in the evening, rather than random arrival and departure at any time during the day. Such occurrences of movement associated with time is common for a building with specific function and an occupant with specific career or post, such as in residential buildings, office buildings, etc., which may be called typical movement patterns for such types of buildings and occupants. An *event* mechanism is proposed to represent and manage the time-triggered occurrences of movement in buildings, in which the movement of occupant is driven by a number of events. In addition, the *event* mechanism can be also used to treat the relevance of the movements of occupants (e.g., a joint movement such as attending a meeting).

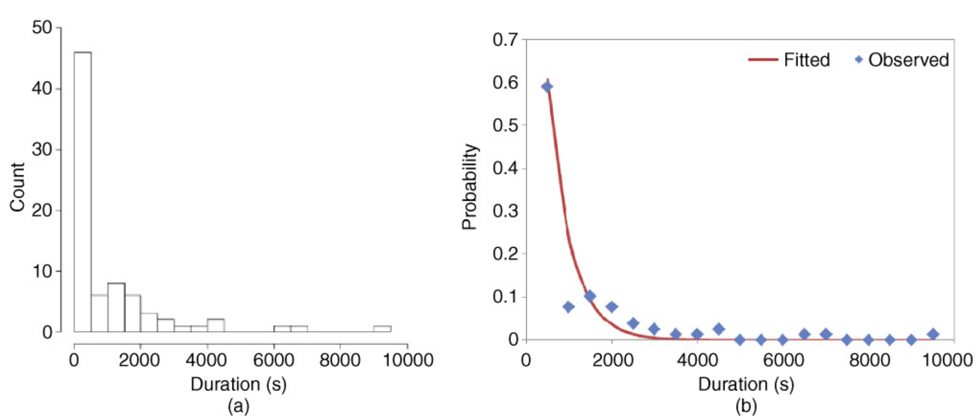


Fig. 4 (a) Frequency distribution of duration of presence, both offices aggregated (78 sojourns of presence, Dodier et al. 2006); (b) fitted and observed probabilities of the presence durations

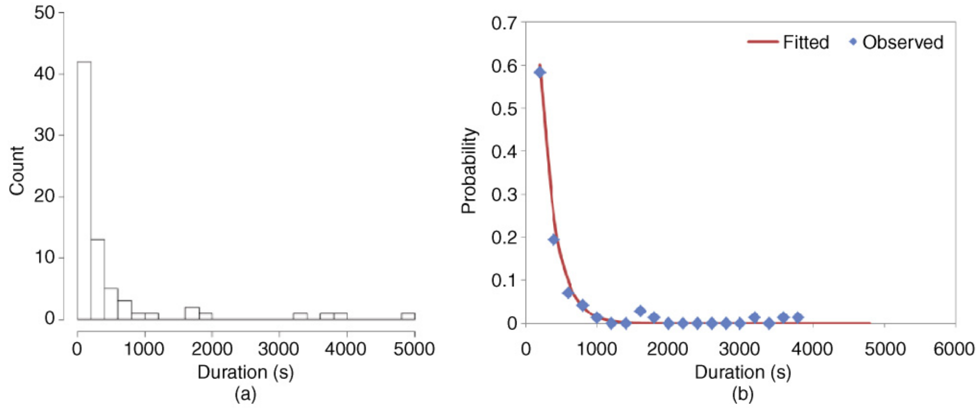


Fig. 5 (a) Frequency distribution of duration of absence, both offices aggregated (72 sojourns of absence, Dodier et al. 2006); (b) fitted and observed probabilities of the absence durations

An event in the model is an object that corresponds to the specific location change of occupants; for example, the event of walking around, regarded as a basic event in all buildings, corresponds to the location change from one space to another (covering general movements, such as going to the washroom, dropping by another office, etc.). The event of going to the office in the office building corresponds to an occupant's location change from outside (space 0) to the office (a space for his/her own office). Each event has a valid period (with a starting time and an ending time) during which it takes place, and a range of actions on the occupants (i.e., it only influences specific occupants). The event drives occupant movement through ***P*** matrix exactly by specifying the corresponding elements in the ***P*** matrix. The probability elements associated with the event are also fixed and time-independent (called "event dependent"). Such transition probabilities can be determined from some statistical indices of the event (see Section 2.2). Each event also has a priority that determines the order of the event taking effect on ***P*** matrix in case that several events are valid at the same time (i.e., valid periods of events have intersections) and these events have common elements in ***P*** matrix. In this situation, such elements of ***P*** matrix would be specified with the associated probabilities of the event with the highest priority.

In summary, an event object usually has six properties: starting time, ending time, locations (from one space to another), participants (taking part in the event), transition probabilities (driving the movement of participants), and priority (to resolve conflicts).

With the event mechanism, the Markov chain of occupant movement is indeed inhomogeneous since the transition probabilities in ***P*** matrix could be changed at certain times. However, the probabilities are fixed for the remaining periods, so most behaviors of such a Markov chain look like a HMC during every period split by events. Our method is still labeled as HMC.

A set of events can be made in chronological order according to the typical movement patterns of building and occupant, and it is easily scalable to describe other events, such as long absence, meeting, short visit, working at home, working part-time, and even other scheduled events. Once the list of events is made, the model is completely constructed in form.

2.1.3 *Algorithm*

Due to the simple structure, the implementation of the model is relatively simple once the building typology, occupant and movement parameters (e.g., the number of occupants, the accessible spaces for each occupant, the set of events with properties, etc.) are determined.

The model was implemented as a MATLAB script. The location of each occupant was simulated independently based on the inputs related to that occupant. For initial states, the occupant is considered to be absent in the case of office buildings (or present in the case of residential buildings) at the initial time step, 00:00 of January 1st. From then on, the time series of occupant location is generated by the transition probability matrix at each time step, which is specified by valid events.

Figure 6 shows how the algorithm works: (0) initialize the states of all occupants at time step 0; for each time step, (1) update the set of active events at present according to the input set of events and their valid periods; (2) update the ***P*** matrices of all occupants according to the set of active events, the corresponding elements of ***P*** matrix are specified by the active events, and note that the sum of elements in each row of ***P*** matrix should equal 1; (3) for each occupant, determine the current state of occupant according to the previous state and the updated ***P*** matrix (see Fig. 7 for details, where the MATLAB function *rand* generates a pseudorandom

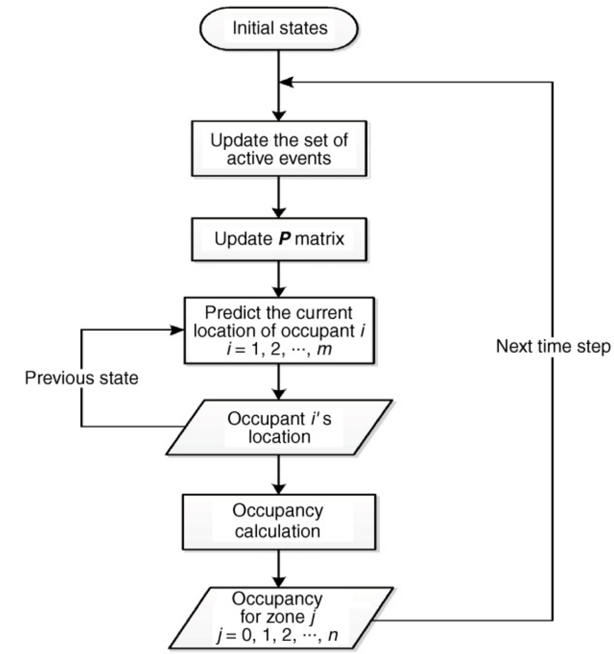


Fig. 6 The workflow of algorithm (processing stage)

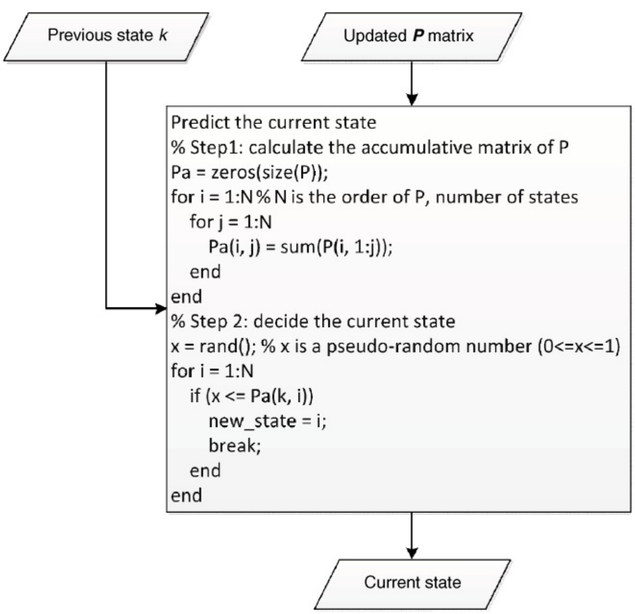


Fig. 7 The simulation script of Markov chain at each time step

value drawn from the standard uniform distribution), and do this for all occupants; (4) calculate the current occupancy for all zones according to the locations of all occupants.

By repeating this procedure step by step, the time series of the location of each occupant and the occupancy of each zone in the building can be generated, without any constraint on the number of occupants and the number of zones.

2.1.4 Reduction and estimation of transition probabilities

The kernel parameters for the model are the transition probabilities for each occupant. Since the probabilities in the model are time-independent, it greatly reduces the complexity in the time dimension compared to the pre-existing models. Actually, all the entries of ***P*** matrix for each occupant can be estimated directly from the information that is collected by deploying sensors in the building which track each individual over time as proposed by Tabak (2008). However, it still seems not trivial in simulation to directly input all the entries of ***P*** matrices of all occupants in multi-occupant multi-zone scenarios as mentioned by Liao et al. (2011). This is due to the high order of matrices (corresponding to the number of spaces) and the number of matrices (corresponding to the number of occupants). Therefore, a way to further simplify the specifications for such matrices is proposed in this paper.

The most important requirement for the simplified method is the generated ***P*** matrix should capture the specific statistical characteristics of building occupancy (i.e., occupant movement). An office building is used as an example to demonstrate the procedure how to apply the present occupancy model to a specific type of building and how to simplify the inputs based on the key statistical characteristic of occupant movement.

2.2 Occupancy modeling in an office building

An office building is a common building type which most occupancy modeling research is focused on. In an office building, a typical movement pattern for the occupants in a workday is going to the office in the morning, having a lunch break at noon, and getting off work in the evening or night (possibly with overtime), which leads to the phased variation of the number of occupants of a typical office building over the workday as shown in Fig. 8. During the working periods, the occupancy for each zone of the building would change since the occupants walk around among the spaces inside and outside of building for a variety of reasons.

In addition to the phased evolution, the occupants' comings and goings for the morning arrival (t_1 to t_2), the

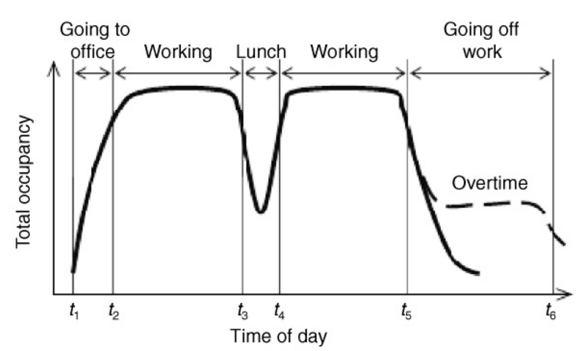


Fig. 8 The occupancy variation of an office building over the workday

lunch break (t_3 to t_4), and the night departure (t_5 to t_6) are all random rather than deterministic (or strictly according to a schedule), which greatly affects the working time of building devices. The indices representing the morning arrival and the night departure are respectively the time of morning arrival and the time of night departure. Those for a lunch break are the time of going out for lunch and the time of coming back to office (it is assumed that the occupants do not have lunch in their own office).

Such a typical movement process of the occupant in an office building can be described by a set of five events: walking around, going to the office, going for lunch, coming back from lunch, and getting off work. Next, an approach to model such events with a specific statistical characteristic and the treatment for events that are not detailed in this paper are discussed.

2.2.1 Walking around

The event of walking around in office buildings corresponds to the general location change from one space to another, among the spaces inside and outside of building (e.g., going to the washroom, walking in the hallway, going outside, dropping by another office, etc.). In most situations, it means a transient movement of an occupant out of his office. Its valid period is usually the same as the business time of a company, e.g., from 8:00 to 17:00.

As discussed above, specifying the $\boldsymbol{P}$ matrix for walking around during the working period is the most challenging work for multi-zone scenarios. Since the present stochastic model should reproduce the specific statistical characteristics of occupant movement, a simplified way can be proposed to determine the probabilities.

From experience, the key statistical characteristic for an occupant walking around is the long-run proportion of time and the expected sojourn time that the occupant stays in each space of the building during the working period. The simplified way to specify $\boldsymbol{P}$ matrix is based on the following mathematical formulations.

<u>The stationary distribution</u>

Let $\boldsymbol{P}$ denote the transition probability matrix that corresponds to the process of an occupant moving among the spaces (all zones in the building and outside), which are indexed by $\{0 = \text{outside}, 1 = \text{zone } 1, 2 = \text{zone } 2, \cdots, n = \text{zone } n\}$. Suppose the probabilities are fixed and not influenced by any other events in normal conditions.

$$\boldsymbol{P} = (p_{ij})_{(n+1)\times(n+1)} = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0n} \\ p_{10} & p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & & \vdots \\ p_{n0} & p_{n1} & \cdots & p_{nn} \end{pmatrix} \tag{2}$$

In our situations, this Markov chain is irreducible and ergodic. Such a Markov chain has a stationary distribution, denoted by $\boldsymbol{\pi}$, $\boldsymbol{\pi} = (\pi_0, \pi_1, \cdots, \pi_n)$, where π_i means the long-run proportion of time that the Markov chain is in state i (i.e., with what proportion of time the occupant stays in the space i). And we have (Ross 1996):

$$\sum_{i=0}^{n} \pi_i = 1 \tag{3}$$

$$\boldsymbol{\pi} = \boldsymbol{\pi P} \tag{4}$$

Based on Eqs. (3) and (4), given the matrix $\boldsymbol{P}$, the vector $\boldsymbol{\pi}$ can be determined by solving Eqs. (3) and (4) simultaneously, and expressed as Eq. (5):

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ p_{01} & p_{11} & \cdots & p_{n1} \\ \vdots & \vdots & & \vdots \\ p_{0n} & p_{1n} & \cdots & p_{nn} \end{pmatrix} \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \boldsymbol{\pi}^{\mathrm{T}} = \boldsymbol{A}^{-1}\boldsymbol{b} \tag{5}$$

where $\boldsymbol{\pi}^{\mathrm{T}}$ is the transpose of $\boldsymbol{\pi}$.

<u>The expected sojourn time</u>

The sojourn time in state i of Markov chain, denoted by ST_i, is geometrically distributed as expressed in Eq. (6). (Instead, it is exponentially distributed in the continuous-time Markov chain.)

$$\boldsymbol{P}\{ST_i = k\} = p_{ii}^{\,k-1}(1 - p_{ii}) \tag{6}$$

where k is the number of time steps during which the Markov chain is out of state i (measured from the time when the Markov chain is in state i).

The expected sojourn time can be expressed by (Sheng et al. 2008):

$$E(ST_i) = \sum_{k=1}^{\infty} k \cdot \boldsymbol{P}\{ST_i = k\} = \sum_{k=1}^{\infty} k \cdot p_{ii}^{\,k-1}(1 - p_{ii}) = \frac{1}{1 - p_{ii}} \tag{7}$$

So that,

$$p_{ii} = 1 - \frac{1}{E(ST_i)} \tag{8}$$

Denote the vector of expected sojourn time by $\boldsymbol{Est}$, $\boldsymbol{Est} = (Est_0, Est_1, \cdots, Est_n)$, where Est_i means the expected sojourn time that the Markov chain is in state i (i.e. how long the occupant stays in the space i at a time).

Equations (5) and (8) illustrate the relations between the $\boldsymbol{P}$ matrix and the long-run proportion of time and the expected sojourn time.

Given the long-run proportion of time $\boldsymbol{\pi}$ and the expected sojourn time ***Est*** of the occupant in every space, the simplified way to specify $\boldsymbol{P}$ matrix can be described as an optimization problem:

$$\begin{aligned} &\min \ \|(\boldsymbol{A}^{-1}\boldsymbol{b})^{\mathrm{T}} - \boldsymbol{\pi}\|_2 \\ &\text{s.t.} \quad p_{ij} \geqslant 0, \ \sum_j p_{ij} = 1 \\ &\qquad p_{ii} = 1 - \frac{1}{Est_i} \end{aligned} \tag{9}$$

where $\tilde{\boldsymbol{\pi}} = (\boldsymbol{A}^{-1}\boldsymbol{b})^{\mathrm{T}}$ is an estimation of $\boldsymbol{\pi}$. This optimization problem can be solved by using the *fmincond* function in MATLAB.

Thus, the $\boldsymbol{P}$ matrix with $(n+1)\times(n+1)$ entries can be specified by using a set of $2\times(n+1)$ parameters, which greatly reduces the space complexity of such inputs. Naturally, the Markov chain governed by this $\boldsymbol{P}$ matrix will have the same statistical characteristic of occupant movement that is defined by the long-run proportions of time and the expected sojourn times of the occupant in every space. These two vectors of an occupant are then used as the input of the present model, denoted by "movement vectors" in the sequel.

By using the movement vectors, the input information for the model is easier to collect by deploying tracking sensors or conducting a questionnaire survey of the occupants' behavior. The price of such simplification is that the produced $\boldsymbol{P}$ matrix might lose some inherent information that relates the movement of occupant in the building compared to a directly specified $\boldsymbol{P}$ matrix, since every element of the $\boldsymbol{P}$ matrix has its own meaning. Whether to choose the simplified method or specify the $\boldsymbol{P}$ matrix depends on the user's demand. In the most situations of building occupancy simulation, the detailed specification of $\boldsymbol{P}$ matrix is time-consuming and not necessary; instead, the movement vectors of occupant are simple, effective, and accurate enough for simulation.

2.2.2 *Going to the office and getting off work*

The event of going to the office (i.e., morning arrival) of an occupant is regarded as the location change from outside the building to his own office. Thus it only relates the elements of the occupant's transition matrix where the row and the column correspond to outside and his office. The valid period of this event is usually a time span before the business hours, e.g. from 7:00 to 8:30, corresponding to the earliest and latest morning arrival time of office workers.

The morning arrival process can be expressed by a two-state HMC with an absorbing state, which is governed by a fixed 2-by-2 $\boldsymbol{P}$ matrix in Eq. (10).

$$P_{\text{go\_office}} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} p_{00} & p_{01} \\ 0 & 1 \end{bmatrix} \end{array} \tag{10}$$

where 0, 1 are respectively the indices of outside and the occupant's office that is known as an absorbing state. The absorbing state means the occupant will definitely enter his office at a certain time, and it means the arrival time of the occupant entering his office when the HMC is in the absorbing state. During the valid period of the event "going to office", the elements of the occupant's $\boldsymbol{P}$ matrix corresponding to outside and his office would be specified with the probabilities in Eq. (10).

Measured from the start of the event of going to the office, the morning arrival time of the occupant is geometrically distributed, and the expected arrival time (denoted by $E(FA)$) can be expressed by

$$E(FA) = \frac{1}{1-p_{00}} \quad p_{00} = 1 - \frac{1}{E(FA)} \tag{11}$$

If the arrival time is the same as the on-duty time, $p_{00} = 0$ and $E(FA) = 1$.

Similarly, the event of getting off work (i.e., night departure) of the occupant corresponds to the location change from his own office to outside the building. Its valid period is usually a time span after the business hours, e.g., 17:00 − 21:00, corresponding to the earliest and latest night departure time of office workers.

The night departure process can be expressed by a two-state HMC governed by the fixed 2-by-2 $\boldsymbol{P}$ matrix in Eq. (12).

$$P_{\text{off\_work}} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 1 & 0 \\ p_{10} & p_{11} \end{bmatrix} \end{array} \tag{12}$$

where 0, 1 are respectively the indices of outside and the occupant's office. Here, outside the building is the absorbing state, which means the occupant will definitely leave his office at a certain time. And it means the departure time of the occupant leaving his office when the HMC is in the absorbing state.

Measured from the start of the event of getting off work, the night departure time of the occupant is also geometrically distributed, and the expected departure time (denoted by $E(LD)$) can be expressed by

$$E(LD) = \frac{1}{1-p_{11}} \quad p_{11} = 1 - \frac{1}{E(LD)} \tag{13}$$

If the departure time is the same as the off-duty time, $p_{11} = 0$ and $E(LD) = 1$.

2.2.3 Lunch break

The lunch break of an occupant can be divided into two events. One is the event of going for lunch, corresponding to the start of lunch break; the other is the event of coming back from lunch, corresponding to the end of lunch break. We suppose the location for the occupant's lunch is out of his office. The two events can be treated similarly like the event of going to office and getting off work.

The process of going for lunch can be expressed by the transition matrix in Eq. (14). Measured with the start of the event of going for lunch (i.e., the earliest leaving time), the expected leaving time for lunch (denoted by ($E(LL)$) is expressed by Eq. (15).

$$P_{\text{lunch\_out}} = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 1 & 0 \\ p_{10} & p_{11} \end{bmatrix} \end{array} \tag{14}$$

$$E(LL) = \frac{1}{1-p_{11}} \quad p_{11} = 1 - \frac{1}{E(LL)} \tag{15}$$

The process of coming back from lunch can be expressed by the transition matrix in Eq. (16). Measured with the start of the event of coming back for lunch (i.e., the earliest return time), the expected return time from lunch (denoted by $E(LR)$) is expressed by Eq. (17).

$$P_{\text{lunch\_back}} = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} p_{00} & p_{01} \\ 0 & 1 \end{bmatrix} \end{array} \tag{16}$$

$$E(LR) = \frac{1}{1-p_{00}} \quad p_{00} = 1 - \frac{1}{E(LR)} \tag{17}$$

2.2.4 Other issues

There are many other things in an office building that have not been discussed in detail in the above paragraphs, such as meetings that happen during a workday, long period absence or leaves or working at home that result in the occupants not entering the office for a whole workday, short visits or working part-time that the visitors or occupants only appear in the building for a short period of workday, tea breaks for some organizations, or the situation that the occupant has his lunch in his office, etc.

The reasons that those situations are not modeled are, on one hand, this paper focuses on the modeling of most typical movement patterns in an office building, it's not necessary to involve too many events; on the other hand, on the basis of our present model, such things can be modeled easily with an external event generator (random or scheduled), whose outputs can be used as the inputs for the model (see Fig. 9). The interactions of occupants, such as going for lunch together, can also be defined in the form of events.

As for the maximum occupancy limits of spaces, a simple rule can be added into the determination of occupant location at each time step, that is, if the space that the occupant would go to has reached its maximum occupancy limit, the occupant would return to his last location rather than enter the zone, i.e., the occupant's location at the current time step would be the same as the previous time step.

2.2.5 Summary

With the events modeled above, a typical movement pattern of occupants over the workday in an office building, i.e., "morning arrival—walking around (working period)— lunch break—walking around (working period)—night departure", can be simulated by a unique Markov chain, in which the transition matrix for each occupant is successively specified with the associated probabilities of events during different periods of the day. Such probabilities for morning arrival, walking around, lunch break, night departure can be determined based on the statistical indices of each event. The simulated results of such a Markov chain will have the same statistical characteristic of events.

Besides the basic information of building and occupant (building typology, occupant number, working schedule, etc.), the particular inputs for the present occupancy model of an office building are clarified as follows.

(1) For walking around, the vectors of long-run proportion of time and expected sojourn time for each space are needed.

(2) For morning arrival, the earliest, the latest and the expected arrival time are needed. For night departure, the earliest, the latest and the expected departure time are needed.

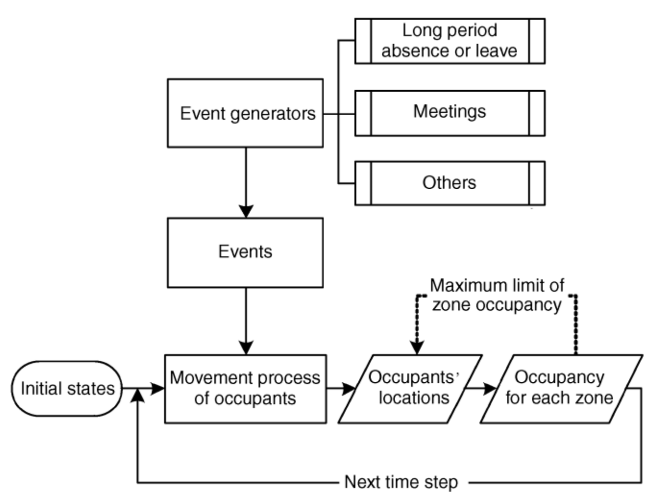


Fig. 9 The integration of event generators within the present model

(3) For lunch break, the earliest, the latest and the expected leaving time for going for lunch, the earliest, the latest and the expected return time for coming back from lunch are needed.

In addition, due to different personalities of occupants, the associated events (or event properties) for each occupant can be different and thus need to be specified individually.

All the above information needs to be collected and calibrated before applying the model.

2.3 Discussion on model calibration and validation

The validation of the proposed stochastic occupancy model, theoretically speaking, should be based on probability, which means the comparison of the probability distribution function (PDF) of the measured and simulated parameters. This requires a long period (usually several years) of measured occupancy data in the real office buildings. But such data are too scarce and not available for most buildings. So the theoretical probabilistic test cannot be carried out due to too few samples. Thus, from the practical point of view, a simple test approach needs to be proposed. For example, the maximum, minimum and average occupancy for each zone inside the building could be chosen as the test parameters and the criterions to calibrate the model. This will be an important work and needs more study in the future.

For the moment, due to the lack of measured data, the proposed model has not been fully validated by comparing the simulated and the measured data. Nonetheless, an illustrative case study can be made to check and test the capacity of the model; whether it can cover the things that affect the occupancy variation in a building and how much it can capture this.

3 Case study

A simple office building is tested to demonstrate the usage and effect of the proposed model. This case is illustrative and the input data are taken from experience. The time step used in the case is 5 min; an occupancy time series of one day is comprised of 288 points.

Through the case, we will check the capacity of the model to represent: (1) the trend of "going to the office—working—lunch break—working—getting off work" in a typical workday; (2) the random arrival time and the smooth increase of building occupancy in the morning; (3) the random departure time (overtime) and the smooth decrease of building occupancy; (4) the random decrease of building occupancy during the lunch break; (5) random movement among the inside and outside spaces during working time.

3.1 Input settings

3.1.1 *Building typology*

The 2D plan of the office building is shown in Fig. 10. There are 4 office rooms, 1 corridor, and 1 restroom, indexed from 1 to 6; and the outside is indexed by 0. There are 7 spaces in total. All spaces are connected by doors and corridor.

3.1.2 *Occupant and movement parameters*

There are 3 types of occupants in the building. The occupants in offices 1 and 2 are ordinary workers. The occupants in office 3 are administrative staff (secretary). The occupant in office 4 is a manager (head of the organization). They move among the 7 spaces (i.e., all spaces are accessible for each occupant). The number of occupants for each office is shown in Table 1.

A fixed working schedule from 8:00 to17:00 is specified. Five typical types of events and a scheduled meeting are considered in the workday. The schedule and events in the workday is shown in Table 2.

The exceeding periods of going to office and returning from lunch, compared to the standard schedule, i.e., 7:00 – 8:00, 8:00 – 8:30 and 13:00 – 13:30 respectively means the early morning arrival, the late morning arrival and the late back to the office from lunch. The exceeding period of getting off work, 17:00 – 21:00, means the delay or overtime for night departure.

The events (except meeting) work for all occupants. The expected values for the events are shown in Table 3. Note that these values are accounted in the unit of 5-minute time step. They are only estimations for the statistical index of each event, and used to determine the transition probabilities in ***P*** matrix and the variations of occupancy ultimately.

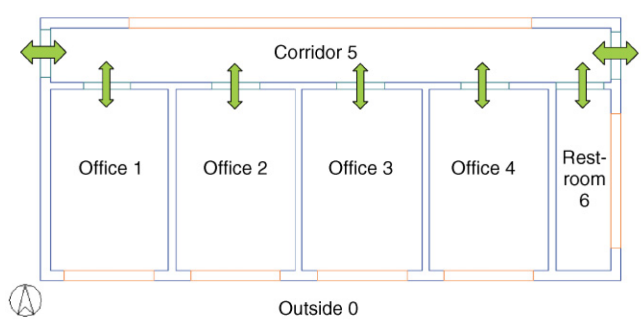


Fig. 10 2D plan of the office building

Table 1 Number of occupants in each office

Room No.	Number of occupants	Room No.	Number of occupants
Office 1	6	Office 2	6
Office 3	2	Office 4	1

Table 2 Schedule and events in a working day

Schedule	Event	Valid period
Working time 8:00–17:00 Lunch break 12:00–13:00	Go to the office	07:00 – 08:30
	Leave for lunch	12:00 – 12:30
	Return from lunch	12:30 – 13:30
	Get off work	17:00 – 21:00
	Walk around	08:00 – 12:00, 13:00 – 17:00
	Meeting	10:00 – 11:30

The expected value of morning arrival time for going to the office is 7:45; which means, in the 5 min unit the geometric distribution of morning arrival time has an expected value of 9 time steps (measured from 7:00).

The values of long-run proportion and mean sojourn time for different spaces mean: for the occupants in office 1, each of them will spend 90% of business hours in his own office, 3% in the other three offices, 1% in the outside, 1% in the corridor, and 5% in the restroom; the mean sojourn time steps for each space are 24, 3, 2, 2, 2.

With different movement vectors, the differences between the three types of occupants are specified. A manager may have more meetings out of the office, so he/she only spends 60% of business hours in the office and 30% outside, with 1-hour mean sojourn in both spaces. A secretary may tend to move more frequently in the spaces, so he/she only spends 70% of business hours in the office and the remaining 30% occupying other spaces, with 1-hour mean sojourn in the office and 10 min for each other space.

To illustrate the effect of meeting events, a scheduled meeting is held in the manager's office from 10:00 to 11:30, and it is attended by two occupants from office 1, two occupants from office 2 and the manager. All the participants are determinate.

3.2 Transition matrix

The transition matrix for each occupant can be determined by the movement parameters according to the equations in Section 2.2. The associated probabilities of events for occupants in office 1 are shown in Fig. 11. Based on those probabilities, the transition matrix would be specified and changed at certain time steps when the events become valid over a workday.

3.3 Results and discussions

The simulation for a workday runs 1000 times consecutively, with different random seeds for each simulation. Figure 12 shows the generated time series of the locations of four occupants in offices 1, 2, 3 and 4. As expected, the four occupants stay in their own offices for most of the time, with different times of morning arrival and night departure, different times of leaving and returning during lunch break, and occasionally out of the office (into other spaces). The scheduled meeting can be seen from that the occupant in office 1 and the manager in office 4 stay in the office 4 during the period of 10:00 – 11:30. The secretary in office 3 seems to move more frequently than others.

3.3.1 Building occupancy

The change of number of occupants in the whole building over a workday is illustrated in Fig. 13, where (a) shows the result of building occupancy for one simulation, (b) shows the results for five successive simulations, (c) shows the maximum, minimum, and average occupancy for each time step from all the simulations, (d) shows the hourly occupancy taking the mean of five-minute results. It can be seen that (1) the trend of "going to the office – working – lunch break – working – going off work" in a typical workday

Table 3 Expected values for each occupant

Event	Statistical index		Expected value
Go to the office	Morning arrival time		7:45
Go for lunch	Leaving time		12:10
Come back from lunch	Return time		12:50
Get off work	Night departure time		18:00
Walk around	Long-run proportion of time and mean sojourn time in each room	Office 1	$\boldsymbol{\pi}$ = [0.01, 0.9, 0.01, 0.01, 0.01, 0.01, 0.05] ***Est*** = [10min, 2h, 15min, 15min, 15min, 10min, 10min]
		Office 2	$\boldsymbol{\pi}$ = [0.01, 0.01, 0.9, 0.01, 0.01, 0.01, 0.05] ***Est*** = [10min, 15min, 2h,15min, 15min, 10min, 10min]
		Office 3	$\boldsymbol{\pi}$ = [0.05, 0.05, 0.05, 0.7, 0.05, 0.05, 0.05] ***Est*** = [10min, 10min, 10min, 1h, 10min, 10min, 10min]
		Office 4	$\boldsymbol{\pi}$ = [0.3, 0.01, 0.01, 0.01, 0.6, 0.02, 0.05] ***Est*** = [1h, 5min, 5min, 5min, 1h, 10min, 10min]

$$
\boldsymbol{P}=\begin{array}{c} \\ 0\\1\\2\\3\\4\\5\\6 \end{array}\!\!\begin{array}{c} \begin{array}{ccccccc}0&1&2&3&4&5&6\end{array}\\ \left[\begin{array}{ccccccc}1&0&0&0&0&0&0\\1&0&0&0&0&0&0\\1&0&0&0&0&0&0\\1&0&0&0&0&0&0\\1&0&0&0&0&0&0\\1&0&0&0&0&0&0\\1&0&0&0&0&0&0\end{array}\right]\end{array},\quad
\boldsymbol{P}=\begin{array}{c} \\ 0\\1\\2\\3\\4\\5\\6 \end{array}\!\!\begin{array}{c} \begin{array}{ccccccc}0&1&2&3&4&5&6\end{array}\\ \left[\begin{array}{ccccccc}0.5000&0.3724&0.0247&0.0247&0.0250&0.0260&0.0272\\0.0042&0.9583&0.0023&0.0023&0.0023&0.0042&0.0263\\0.0244&0.2120&0.6667&0.0234&0.0237&0.0243&0.0255\\0.0242&0.2131&0.0233&0.6667&0.0235&0.0241&0.0253\\0.0247&0.2110&0.0237&0.0236&0.6667&0.0246&0.0258\\0.0257&0.3747&0.0242&0.0243&0.0245&0.5000&0.0267\\0.0055&0.4734&0.0052&0.0052&0.0053&0.0055&0.5000\end{array}\right]\end{array}
$$

Night time (initial state) Walking around

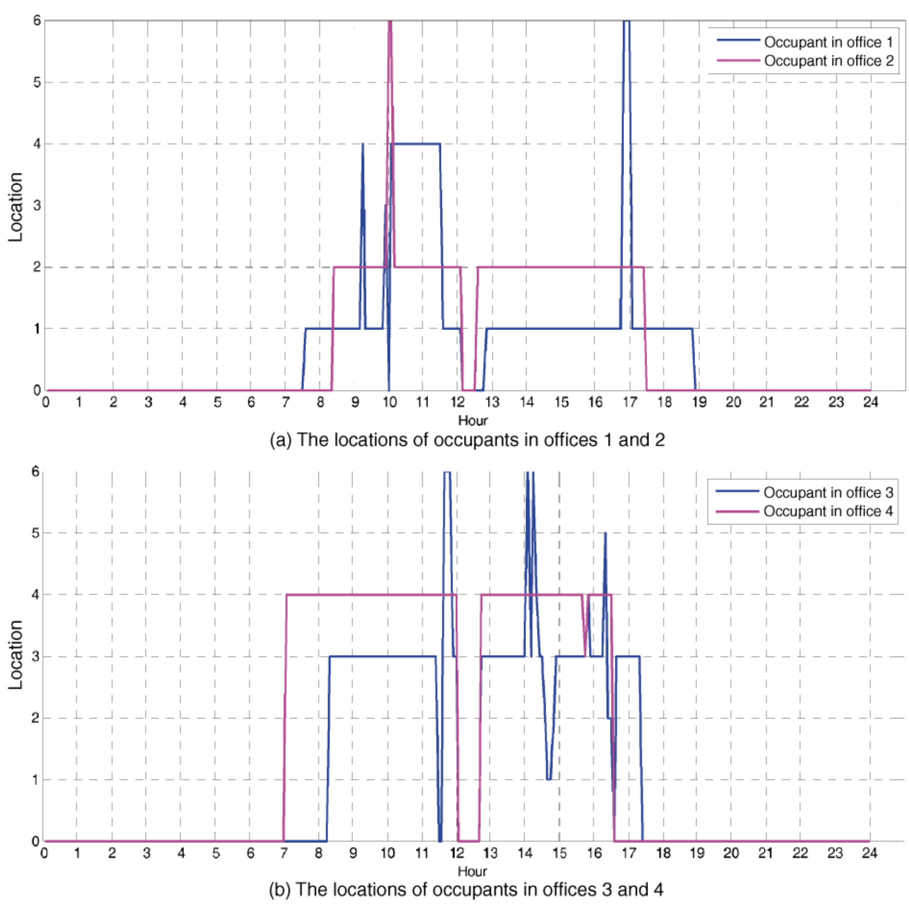
$$
\boldsymbol{P}_{\text{go\_office}}=\begin{array}{c} \\ 0\\1 \end{array}\!\!\begin{array}{c} \begin{array}{cc}0&1\end{array}\\ \left[\begin{array}{cc}0.8889&0.1111\\0&1.0000\end{array}\right]\end{array},\quad
\boldsymbol{P}_{\text{off\_work}}=\begin{array}{c} \\ 0\\1 \end{array}\!\!\begin{array}{c} \begin{array}{cc}0&1\end{array}\\ \left[\begin{array}{cc}1.0000&0\\0.0833&0.9167\end{array}\right]\end{array}
$$

Going to the office Going off work

$$
\boldsymbol{P}_{\text{lunch\_out}}=\begin{array}{c} \\ 0\\1 \end{array}\!\!\begin{array}{c} \begin{array}{cc}0&1\end{array}\\ \left[\begin{array}{cc}1.0000&0\\0.5000&0.5000\end{array}\right]\end{array},\quad
\boldsymbol{P}_{\text{lunch\_back}}=\begin{array}{c} \\ 0\\1 \end{array}\!\!\begin{array}{c} \begin{array}{cc}0&1\end{array}\\ \left[\begin{array}{cc}0.7368&0.2632\\0&1.0000\end{array}\right]\end{array}
$$

Leaving for lunch Returning from lunch

Fig. 11 The transition matrices of occupant in office 1



(a) The locations of occupants in offices 1 and 2

(b) The locations of occupants in offices 3 and 4

Fig. 12 The time series of occupants' locations over a workday

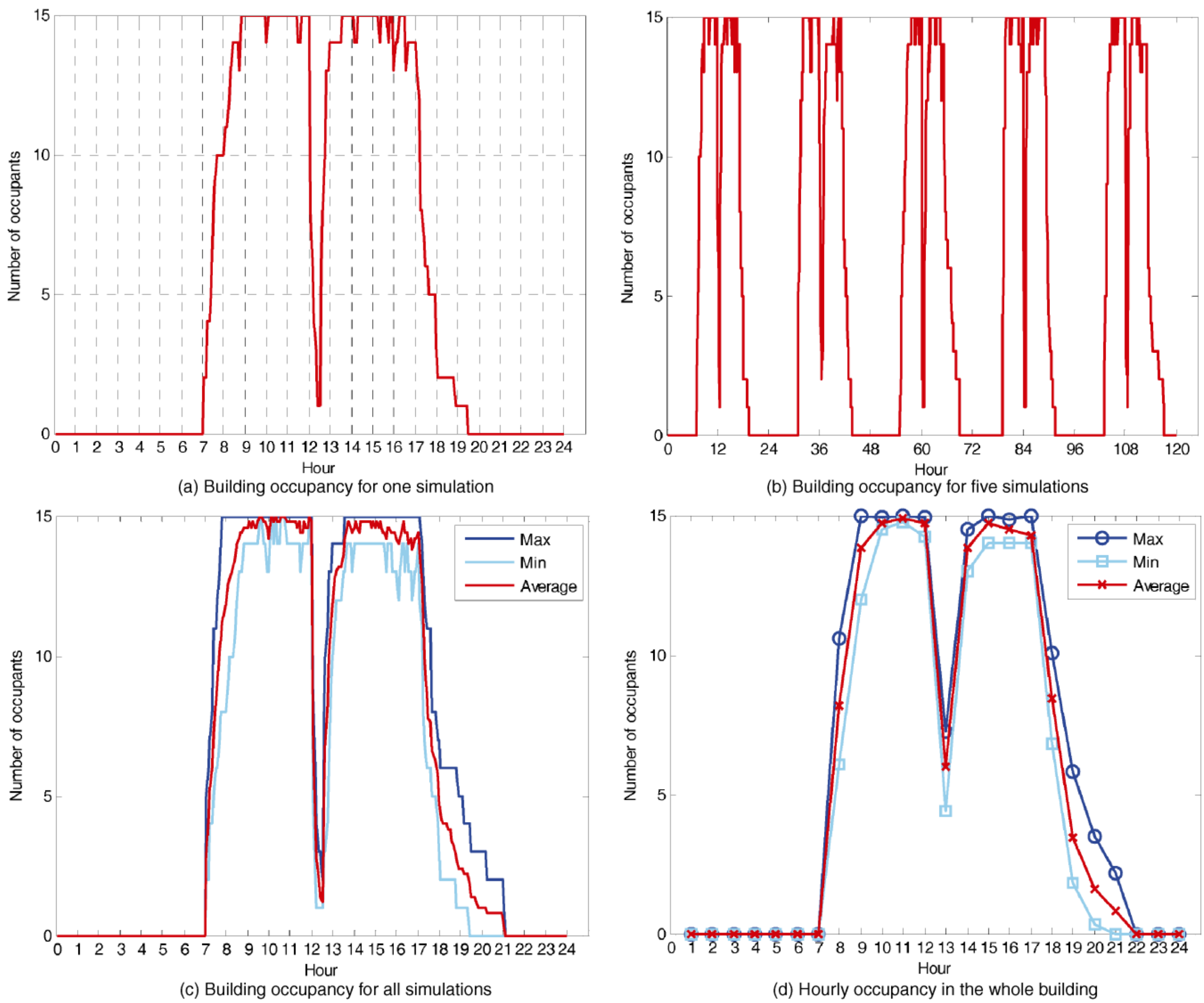


Fig. 13 Building occupancy over a workday

is reproduced; (2) the total building occupancy reaches maximum gradually, rather than sharply under a fixed schedule; (3) during the working time, the total occupancy varies due to the movement of occupants. Such a workday process changes at every simulation, which is understood as random in everyday life.

3.3.2 Zone occupancy

Figures 14, 15, 16, 17, 18 and 19 show the changes of number of occupants in office 1, office 2, office 3, office 4, corridor 5, and restroom 6 over a workday, where (a) is the result of zone occupancy for one simulation, (b) illustrates the maximum, minimum, and average occupancy for each time step from all the simulations, (c) shows the hourly occupancy results. It can be seen that due to the movement of occupants from one space to another, the occupancy of zones inside the building stochastically change in time, more or less than the design value. Although the absence probability for each occupant out of his office (i.e., the probability of staying in other spaces) is as small as 0.1, the resulted occupancy variation is remarkable and should not be neglected in building energy simulation. Due to the scheduled meeting, the occupancy in offices 1, 2, and 4 change a lot during the meeting period.

Since the occupancy model is based on the process of occupant movement, the relationships of stochastic occupancy in multiple zones are realistically taken into account. This results in the reasonable occupancy distribution in space, that is, an increase of occupancy in one space usually means a decrease of occupancy in another space while the

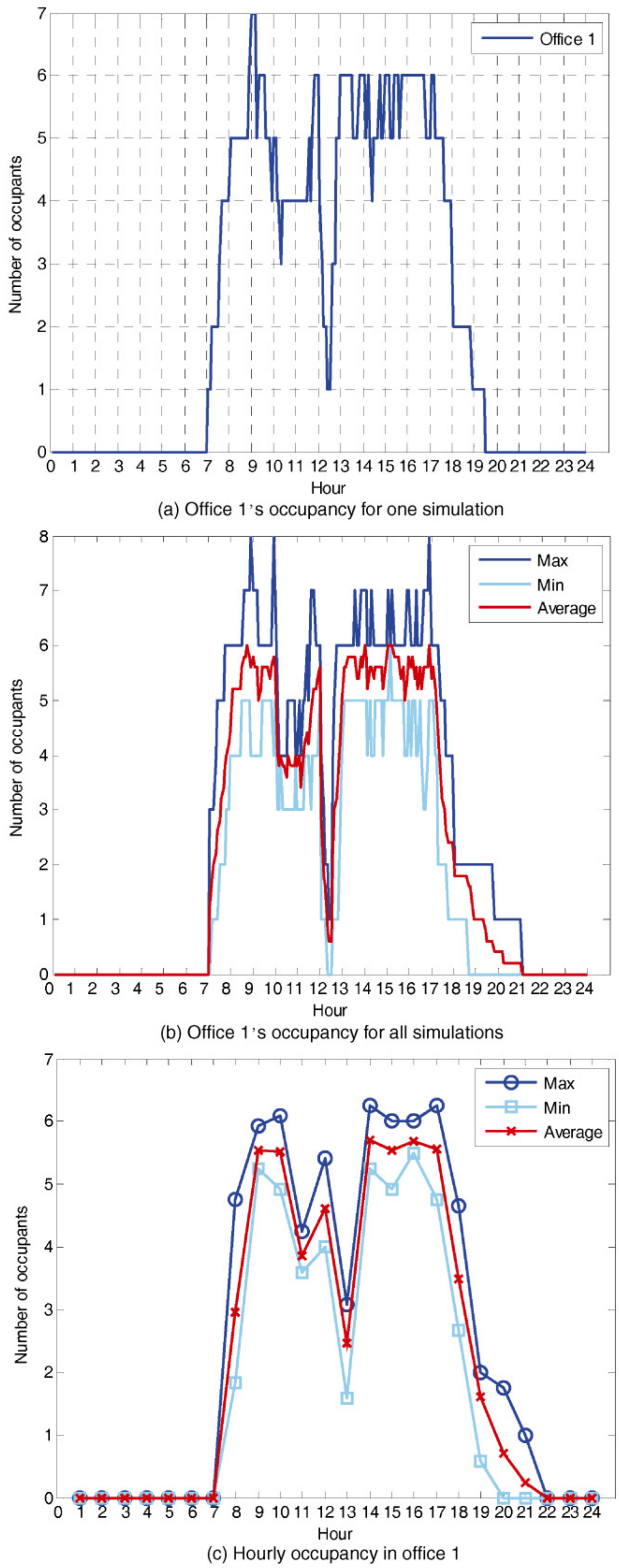


Fig. 14 Office 1's occupancy over a workday

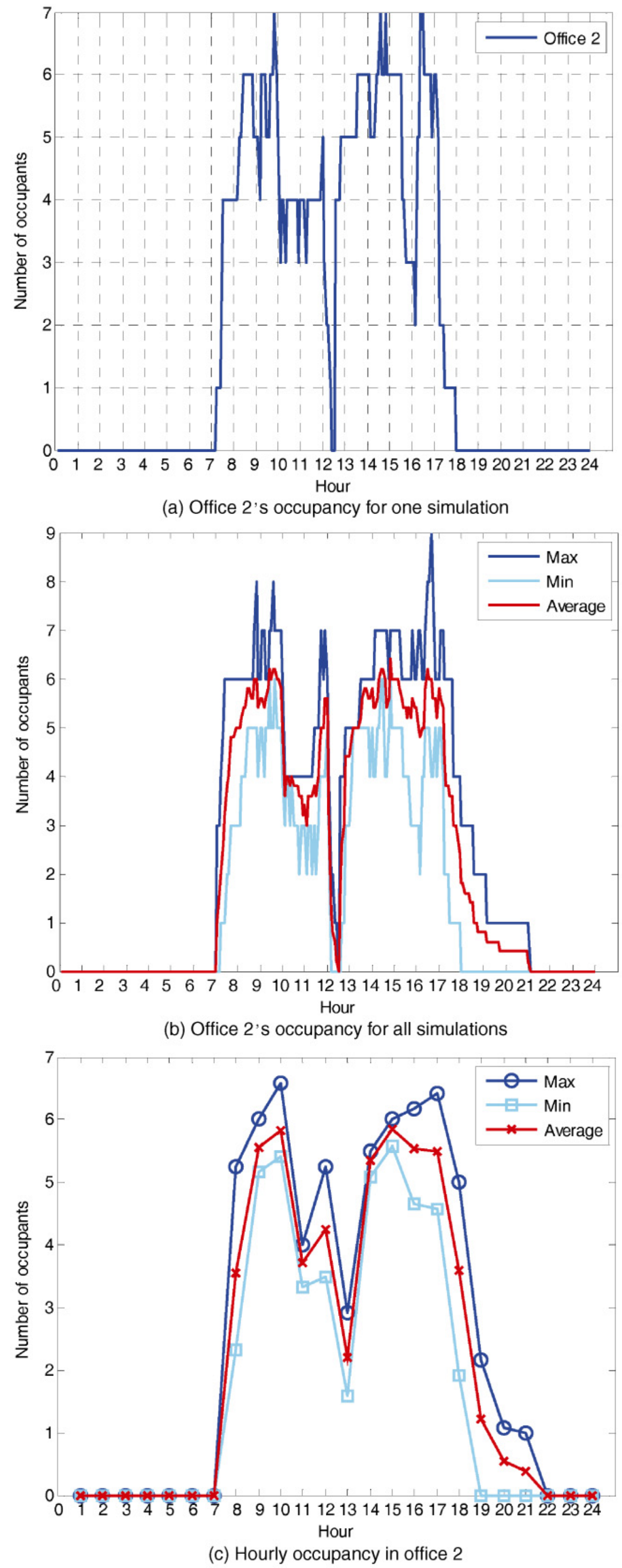
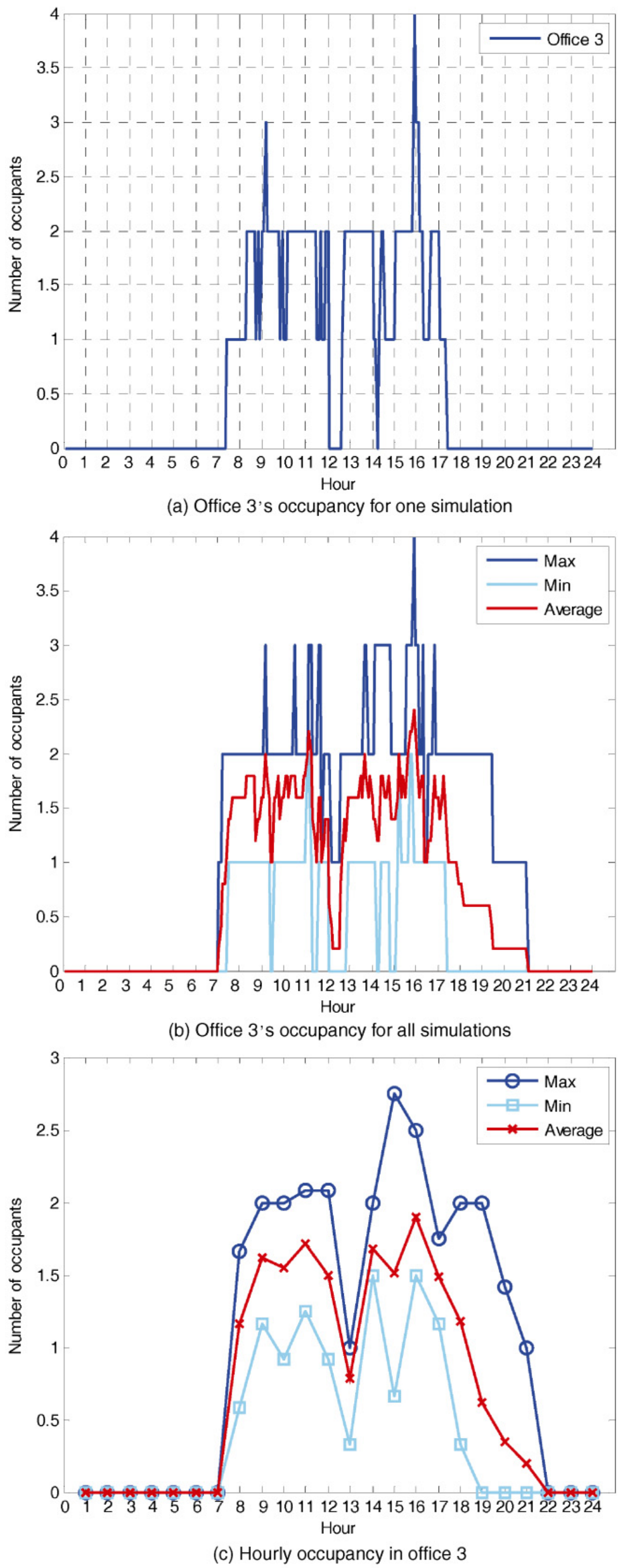


Fig. 15 Office 2's occupancy over a workday

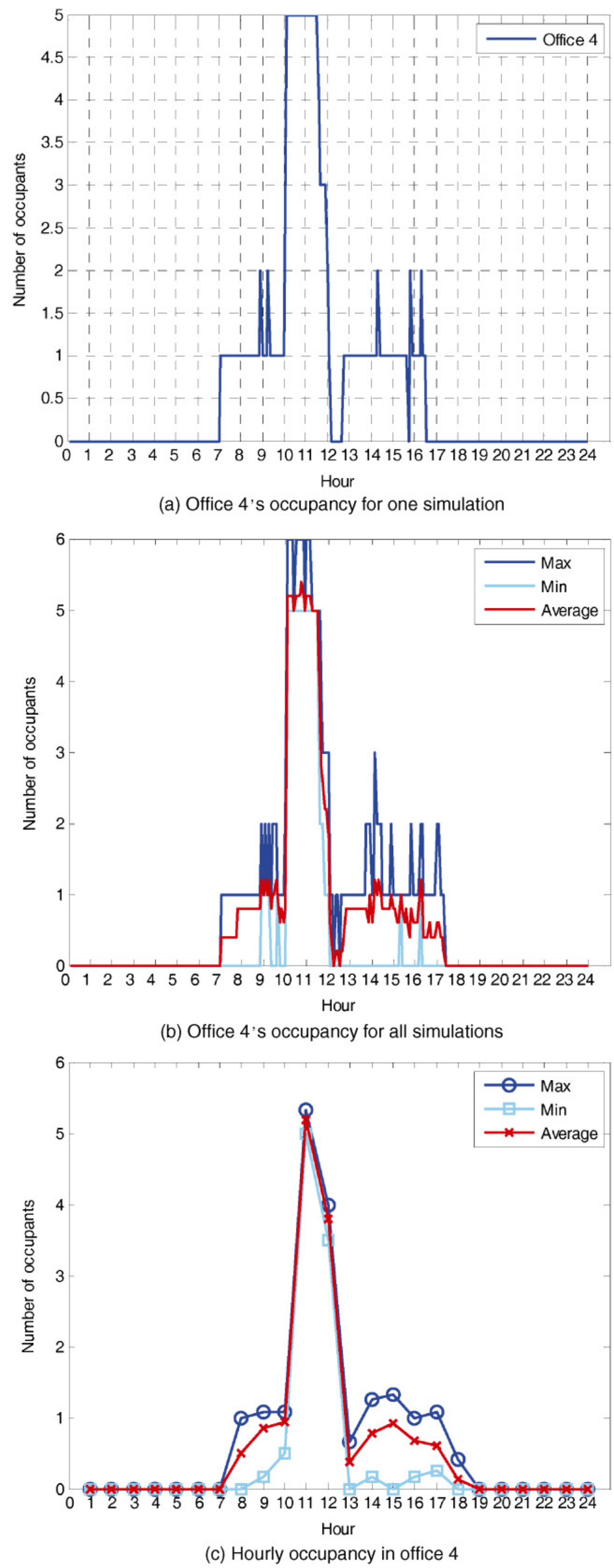


(a) Office 3's occupancy for one simulation

(b) Office 3's occupancy for all simulations

(c) Hourly occupancy in office 3

Fig. 16 Office 3's occupancy over a workday

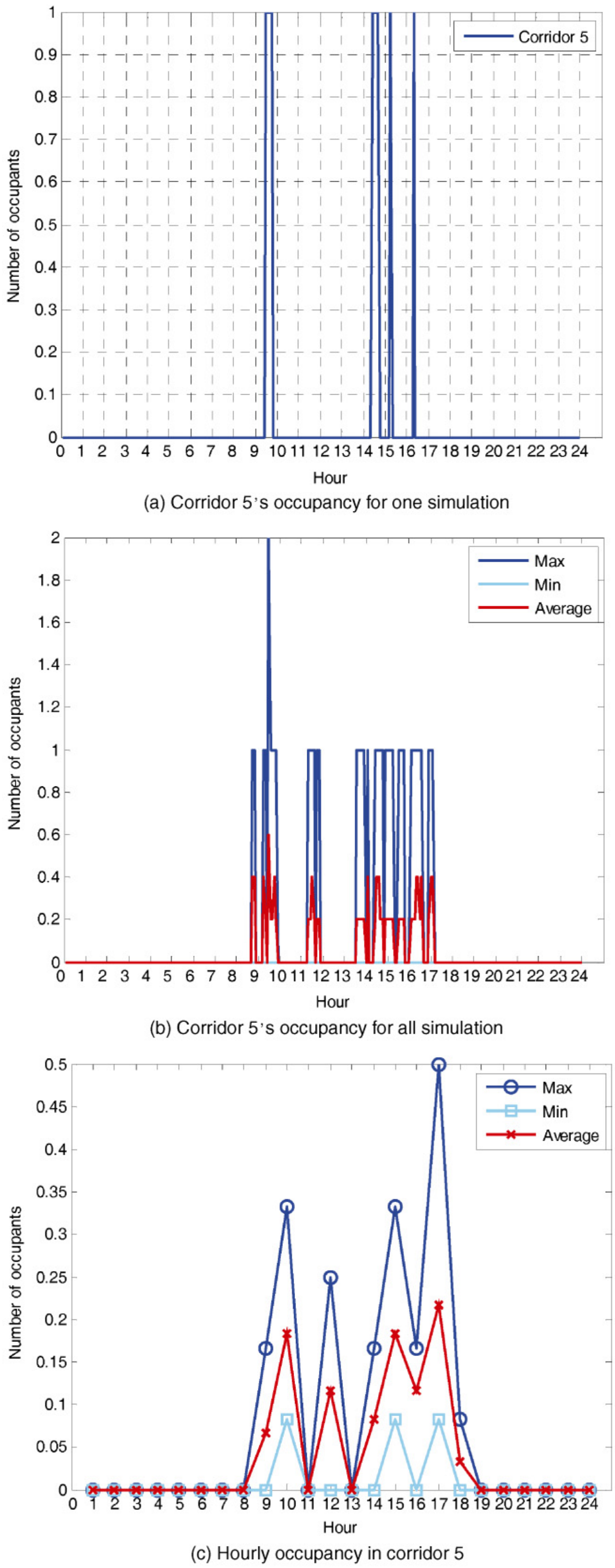


(a) Office 4's occupancy for one simulation

(b) Office 4's occupancy for all simulations

(c) Hourly occupancy in office 4

Fig. 17 Office 4's occupancy over a workday

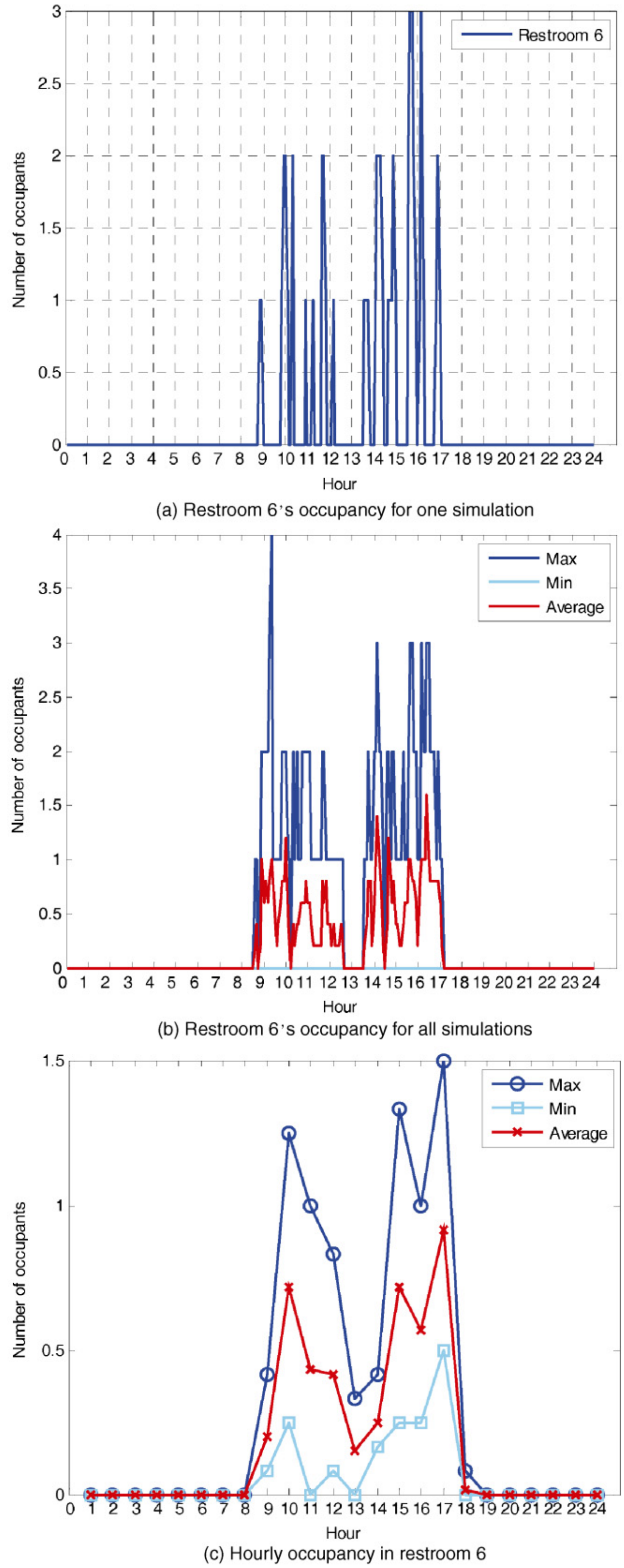


(a) Corridor 5's occupancy for one simulation

(b) Corridor 5's occupancy for all simulation

(c) Hourly occupancy in corridor 5

Fig. 18 Corridor 5's occupancy over a workday



(a) Restroom 6's occupancy for one simulation

(b) Restroom 6's occupancy for all simulations

(c) Hourly occupancy in restroom 6

Fig. 19 Restroom 6's occupancy over a workday total occupancy stays the same. From the simulation results, the occupancy of corridor and restroom are strong coupled with the occupancy of offices and their variations are more significant in time. Even for the spaces with similar function type, the transient occupancy schedules from office 1 to office 4 are not synchronous at each time step (see Fig. 20). Such an uneven distribution of occupancy in space and nonsynchronous change in time would affect the performance evaluation of HVAC systems in simulation.

In general, the stochastic occupancy over a typical workday in an office building can be realistically produced by using the proposed model. Further analysis of simulation results can be made for exploring the validation approach of the model in the future.

4 Conclusions

Building occupancy is a key factor to accurately predict building energy consumption and evaluate the energy saving potential of occupancy-based control system and the performance of HVAC systems. However, it is hard to represent due to its temporal and spatial stochastic nature.

This paper presents a novel approach for occupancy simulation based on the homogeneous Markov chain. In this study, occupancy is handled as the straightforward result of occupant movements among the inside and outside spaces of a building. By using the Markov chain method to simulate the stochastic movement process, the model can generate the location for each occupant and the occupancy for each

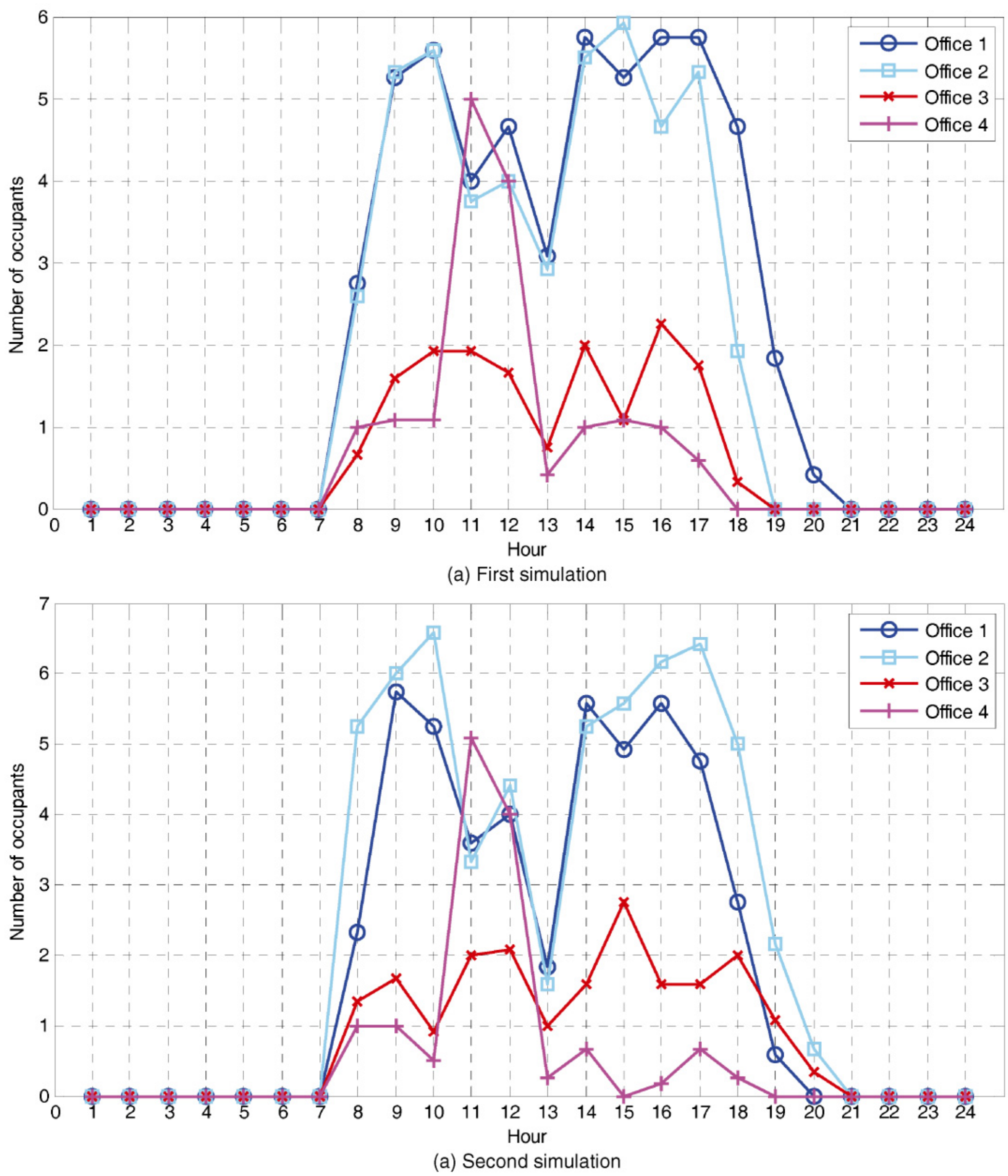


Fig. 20 Four offices' occupancy over a workday

zone of the building. Through this approach, the Markovian property of the state of occupant location is retained, which is validated by other experiments, and the relationships of stochastic occupancy in multiple spaces are realistically taken into account. By using the event mechanism, this model is capable of covering most things that affect the occupancy variation in a building and capturing the movement differences of different types of occupants as well.

From the case study of an office building, it can be seen that the model can produce the realistic occupancy variations in the office building for a typical workday with key statistical properties of occupancy such as the time of morning arrival and night departure, lunch time, periods of intermediate walk-around, etc. Especially, it can produce the nonsynchronous change of occupancy in time and the uneven distribution of occupancy in space, which can distinctly affect the performance evaluation of HVAC systems in a simulation.

The model is simple, clear and has no explicit or implicit constraint with the number of occupants and the number of zones. On the strict mathematical basis of geometric distribution, the model builds the relations of the statistical indices of building occupancy such as mean time of morning arrival and night departure, long-run proportion of time and expected sojourn time. Thus it overcomes the issue of specifying the transition matrix for multi-zone scenarios and maintains a simple, clear set of input parameters. In terms of simplicity, accuracy and unrestraint, this model is sufficient and practical to simulate occupancy for building energy simulations and stochastic analysis of building HVAC systems.

The occupancy model's assumption that the location of an occupant due to movement has a Markovian property is supported by some experiments for single offices but more validations need to be carried out in the future. More events such as short visits should be taken into account in office building to capture the stochastic occupancy variations. From a practical point of view, a simple validation and calibration approach needs to be proposed. The capability of the model in other types of buildings, such as residential buildings, should also be tested and calibrated with specific occupant movement patterns.

Acknowledgements

This project is financially supported by the National Natural Science Foundation of China (No. 51008176), the Specialized Research Fund of the Doctoral Program of Higher Education of China (No. 20100002120014), and Tsinghua University Initiative Scientific Research Program (No. 2009THZ0).

References

Crawley DB, Lawrie LK, Winkelmann FC, Buhl WF, Huang YJ, Pedersen CO, Strand RK, Liesen RJ, Fisher DE, Witte MJ, Glazer J (2001). EnergyPlus: Creating a new-generation building energy simulation program. *Energy and Buildings*, 33: 319 – 331.

Davis JA, Nutter DW (2010). Occupancy diversity factors for common university building types. *Energy and Buildings*, 42: 1543 – 1551.

DeST (2008). DeST-c User Manual. DeST Group, Tsinghua University, China.

Dodier RH, Henze GP, Tiller DK, Guo X (2006). Building occupancy detection through sensor belief networks. *Energy and Buildings*, 38: 1033 – 1043.

EnergyPlus (2009). Input/Ouput Rerference, Version 4.0 Documentation. University of Illinois and Ernest Orlando Lawrence Berkeley National Laboratory, USA.

ESRU (1999). ESP-r: A building and plant energy simulation environment, User Guide Version 9 Series. ESRU Publication, University of Strathclyde, Glasgow.

Goldstein R, Tessier A, Khan A (2010). Schedule-calibrated occupant behavior simulation. In: Proceedings of the Symposium on Simulation for Architecture and Urban, Orlande, USA.

Klein SA, Beckman WA, Mitchell JW, Duffie JA, Duffie NA, Freeman TL, Mitchell JC, Braun JE, Evans BL, Kummer JP, Urban RE, Fiksel A, Thornton JW, Blair NJ, Williams PM, Bradley DE, McDowell TP, Kummert M (2004). TRNSYS 16—A transient system simulation program, User manual. Solar Energy Laboratory, University of Wisconsin-Madison.

Liao C, Lin Y, Barooah P (2011). Agent-based and graphical modelling of building occupancy. *Journal of Building Performance Simulation*, doi: 10.1080/19401493.2010.531143.

Macdonald I, Strachan P (2001). Practical application of uncertainty analysis. *Energy and Buildings*, 33: 219 – 227.

Page J, Robinson D, Morel N, Scartezzini JL (2008). A generalized stochastic model for the simulation of occupant presence. *Energy and Buildings*, 40, 83 – 98.

Ross SM (1996). Stochastic Processes, 2nd Edn. John Wiley & Sons, Inc.

Sheng Z, Xie SQ, Pan CY (2008). Probability and Mathematical Statistics, 4th Edn. Beijing: Higher Education Press. (in Chinese)

Tabak V (2008). User simulation of space utilisation: system for office building usage simulation. PhD Thesis, Eindhoven University of Technology, Netherlands.

Wang D, Federspiel CC, Rubinstein F (2005). Modeling occupancy in single person offices. *Energy and Buildings*, 37: 121 – 126

Yan D (2005). HVAC system simulation and analysis based on uncertain inner heat gains. PhD Thesis, Tsinghua University, China.

Yan D, Xia J, Tang W, Song F, Zhang X, Jiang Y (2008). DeST—An integrated building simulation toolkit, Part I: Fundamentals. *Building Simulation*, 1: 95 – 110.

Zhang X, Xia J, Jiang Z, Huang J, Qin R, Zhang Y, Liu Y, Jiang Y (2008). DeST—An integrated building simulation toolkit, Part II: Applications. *Building Simulation*, 1: 193 – 209