Modelling method of inter-building movement for campus-scale occupancy simulation: A case study

Mingya Zhu¹, Yiqun Pan¹ (✉), Zejun Wu¹, Zhizhong Huang², Risto Kosonen³,⁴

1. School of Mechanical Engineering, Tongji University, Shanghai, China
2. Sino-German College of Applied Sciences, Tongji University, Shanghai, China
3. Department of Mechanical Engineering, Aalto University, Espoo, Finland
4. College of Urban Construction, Nanjing Tech University, Nanjing, China

Abstract
As an important factor in the investigation of building energy consumption, occupant behavior (OB) has been widely studied on the building level. However so far, studies of OB modelling on the district scale remain limited. Indeed, district-scale OB modelling has been facing the challenges from the scarcity of district-scale data, modelling methods, as well as simulation application. This study initiates the extrapolation of occupancy modelling methodology from building level to district scale through proposing modelling methods of inter-building movements. The proposed modelling methods utilize multiple distribution fittings and Bayesian network to upscale the event description methods from inter-zone movement events at the building level to inter-building movement events at the district level. This study provides a framework on the application of the proposed modelling methods for a university campus in the suburbs of Shanghai, taking advantages of data sensing, monitoring and survey techniques. With the collected campus-scale occupancy data, this paper defines five patterns of inter-building movement. One pattern represents the dominated inter-building movement events for one kind of students in their daily campus life. Based on the quantitative descriptions for various inter-building movement events, this study performs the stochastic simulation for the campus district, using Markov chain models. The simulation results are then validated with the campus-scale occupancy measurement data. Furthermore, the impact of inter-building movement modelling methods on building energy demand is evaluated for the library building, taking the deterministic occupancy schedules suggested by current building design standard as a baseline.

Keywords
occupancy modelling; event description; inter-building movement; stochastic process; transition probability; campus buildings; data acquisition

Article History
Received: 16 May 2022
Revised: 30 August 2022
Accepted: 29 September 2022
© Tsinghua University Press 2022

1 Introduction
1.1 OB modelling for building- and district-scale energy simulation

Occupant behavior (OB) modelling has been widely used on the building level, and it has become an important factor in the investigation of building energy consumption (Dong et al. 2021; Tang et al. 2021), building space and facility management (Stjelja et al. 2020), as well building HVAC occupant-centric controls (Yang et al. 2022). Majority of current studies have devoted to the quantification and modelling of occupant behavior within the building space (Dong et al. 2018), for the purpose of understanding how many persons in the space, how people use a space and how their behavior impacts on a building’s energy performance and indoor air quality, together with control of airborne infections risk nowadays (Carlucci et al. 2020; Du et al. 2020; Qian et al. 2020; Wang et al. 2022). Further research often aims to address a breadth of optimization, control and occupant-centric challenges for the operation of individual buildings (Mirakhorli and Dong 2016; Park et al. 2019; Ouf et al. 2021; Zhang et al. 2022). Fueled by rapid improvement of smart sensing, advanced modelling, powerful algorithms of
interpretation and prediction, various occupant behavior modelling methods characterizing complex occupant presence, movement, activities and their determinants (Hou et al. 2020) have been utilized for various applications including building energy performance analysis, building architectural and engineering design, intelligent building operations, building safety design, and real estate business, etc. (Dong et al. 2018; Jin et al. 2021; Zhou et al. 2021; Wang et al. 2022)

This paper has summarized the previous reviews of OB modelling from the perspective of model scale, purpose of review, comparative aspects, and limitations in Table 1. Undoubtedly, recent studies have laid consolidated foundation for further extension of occupant behavior modelling from building level to larger scales, such as district, community, campus and even urban scale. However so far, studies of OB modelling on the district scale remain limited and the impact of OB on energy use should be explored at various temporal and spatial resolutions (An et al. 2017; Yan et al. 2017; Happle et al. 2018). For systematic discrepancy between simulated and measured energy consumption, it has been noted that the poorly estimated occupancy could lead the significant performance gap (Menezes et al. 2012; Brøgger and Wittchen 2018). Similar finding has been discussed in the field of simplified and bottom-up urban building energy models (Kazas et al. 2017; Mastrucci et al. 2017; Mosteiro-Romero et al. 2017). Our review also found two kinds of studies attract growing attention in the field of district energy simulation. One is district energy prediction for retrofit design and evaluation (Orehounig et al. 2014; Chen et al. 2017). The other is understanding the spatiotemporal patterns of district energy demand for system planning and optimization (Yamaguchi and Shimoda 2010; Fonseca and Schluter 2015). Under this context, deep investigations into the effect of OB on the district scale is crucial.

Given the relatively close-off management for university campuses, especially in the Covid-19 period, we believe that the exploration on occupancy modelling methods for campus buildings is a good start for upsaling the current methods. This study attempts to find a good way to describe occupants’ inter-building movements and proposes feasible methods for campus-scale occupancy modelling, taking advantages of data sensing, monitoring and survey techniques.

### 1.2 Challenges of occupancy modelling on the district scale

Even though various modelling approaches have been developed to simulate building occupancy, such as Markov chain (MC) method (Page et al. 2008), logistic regression (LR) (Andersen et al. 2014), agent-based method (ABM) (Liao et al. 2012), and decision tree(s) (DT) (D’Oca and Hong 2015), there has rare application of occupancy models on the district scale. Majority of recent researches have been limited at the room space and building levels due to uncertain modelling method for the upscaled extrapolation

### List of symbols

| Symbol | Description |
|--------|-------------|
| \(K\)  | number of resulted clusters |
| L1-dor | location 1: Dormitory |
| L2-rest | location 2: Restaurant |
| L3-col | location 3: College Lab |
| L4-lib | location 4: Library |
| L5-gym | location 5: Gym |
| L6-lec | location 6: Lecture hall |
| L7-out | location 7: Outside campus |
| \(P(\text{location}_{j},t)\) | probability of locating at \(j\) in \(t\) time |
| \(P(\text{location}_{j},|\text{location}_{i},t)\) | conditional probability of locating at \(i\) in \(t-1\) time and locating at \(j\) in \(t\) time |
| \(P_s\) | Markov chain transition probability matrices |
| \(t_b, t_b'\) | event period of Breakfast arrival \((t_b)\) and Breakfast departure \((t_b')\) |
| \(t_d, t_d'\) | event period of Dinner arrival \((t_d)\) and Dinner departure \((t_d')\) |
| \(t_e\) | event period of Back to dormitory \((t_e)\) |
| \(t_k, t_k'\) | event period of the \(N\)th arrival at lab \((t_k)\) and the \(N\)th departure from lab \((t_k')\) |
| \(t_l, t_l'\) | event period of Lunch arrival \((t_l)\) and Lunch departure \((t_l')\) |
| \(t_o, t_e\) | event period of the \(N\)th arrival at lab \((t_o)\) and the \(N\)th departure from lab \((t_e)\) |

### Abbreviations

- 7-/16-D: 7-/16-dimensional
- ABM: agent-based method
- BAS: building automation system
- DT: decision tree
- GIS: geographic information system
- ICT: information and communication technology
- IEA: International Energy Agency
- LR: logistic regression
- MC/MCM: Markov chain method
- MCMC: Markov chain Monte Carlo simulation
- OB: occupant behavior
and the lack of community/district/urban-scale data. Few studies on university campus occupancy just applied building-level occupancy modelling methods to different campus buildings and distinguished different occupancy patterns on the basis of occupancy profile within a single building (Ding et al. 2019; Ding et al. 2021), rather than directly targets the inter-building movement of different occupants. Indeed, for district scale, it faces several challenges to figure out the occupants’ interaction between buildings under the complex impacts of single building usage and the district-scale environment on the occupants.

According to the definition of OB in IEA Annex 66, occupancy has three features, stochasticity, diversity and complexity (Yan et al. 2017). Although stochasticity has been the research focus over a past decade and has enriched approaches for random occupancy chain modelling in buildings, occupancy is not purely random but could still be characterized by deterministic factors in surrounding systems (Hou et al. 2020). It could be a promising way of understanding and enhancing the methods of characterizing the complexity and diversity of occupancy patterns by incorporating both random and deterministic conditions of occupancy in buildings. The key difference between building-scale and district/urban-scale energy modelling are the various system interactions that should be considered in the context of district/urban environment. The interaction of occupants within/between multiple buildings in the district seems somehow overlooked in district/urban building energy modelling because of the complexity of the larger scale system (Happle et al. 2018). For example, an occupant who is absent from office at lunch time is likely to be present in a restaurant nearby, which is commonly considered in land-use and transport modelling and simulation (Horni et al. 2009).
As a solid methodological foundation for district-scale research, current occupant behavior modeling approaches can be categorized into deterministic space-based approaches, stochastic space-based approaches, and stochastic person-based approaches (Happle et al. 2018). The three previous categories can be further divided into sub-categories according to whether they are able to account for inter-individual diversity among spaces or persons. In this case, deterministic approaches apply schedules of the occupant presence for every hour of a typical day and deterministic rule sets that assume direct causal links between certain drivers and an action on the basis of monitored data (Mahdavi and Tahmasebi 2015). Stochastic approaches focus the correlations between occupant behavior and environmental conditions or the occurrence of specific events, e.g., an occupant’s arrival to the space at the certain time. First-order time-inhomogeneous Markov chain is the commonly used basis of stochastic occupant presence models (Page et al. 2008; Aerts et al. 2014). Their transition probability matrices are usually derived from survey data, describing daily activities of occupant in form of a diary. Stochastic models then typically sample from statistical distributions to predict the likelihood that certain situations or actions occur (Happle et al. 2018). Within the scope of district/urban building energy modelling, deterministic space-based approach is the primary choice due to the fact that they can easily be input to commonly-used building energy simulation tools for building retrofit, district energy system design and operation studies. Based on large survey and behavior questionnaire, several studies have demonstrated the superiority of stochastic person-based approaches over deterministic approaches (He et al. 2015; An et al. 2017), and the effect of OB stochasticity and diversity on district energy demand (He et al. 2015; Baetens and Saelens 2016; An et al. 2017). Compared to deterministic approaches, models considering stochastic occupant behavior are able to generate more realistic spatiotemporal energy demand patterns at the district level (Happle et al. 2018). While, the above studies using stochastic approaches considers limited building types, e.g. residential or offices, or just focused on single buildings at the district-scale.

From the perspective of data basis, exploration of the diversity of occupancy patterns from large data streams will allow for a better understanding of the energy usage in buildings (Nguyen and Aiello 2013). Diversity of occupant behavior patterns becomes obvious when it comes in the contexts of mixed-use district-/urban-scale modelling. It poses challenges on the amount and scale of data acquisition for district-scale occupancy modelling.

In general, most of the existing datasets just cover the occupant presence and movement monitoring within the scale of building or room space. Sensing data of occupant presence and numbers for a certain period in specific room spaces has become an essential basis for occupancy modelling of diverse patterns. Different data sets of the occupancy in buildings have been established for deterministic occupancy schedules learning (D’Oca and Hong 2015) and stochastic occupancy/occupant Markov chain (MC) models development (Wang et al. 2017; Gu et al. 2018). Fueled by rapid development of artificial intelligence, intelligent sensing, information and communication technologies (ICTs), and occupancy detection (Ding et al. 2022), occupancy database establishment could be an irresistible trend for both fundamental and applicant research on occupancy modelling with the consensus cognition that insights of practical data are essential supports for OB modelling methodology development. Several studies have collected actual occupancy data on sites using questionnaires, observations, sensor direct detection (Benezeth et al. 2011; Labeodan et al. 2015), indirect detection (Dong et al. 2010) and mobile device detection (Mohammadi and Taylor 2017; Wang et al. 2017; Gu et al. 2018) methods. In the timeline from Annex 66 to 79 research, OB data collection and the further generation of large-scale datasets have always been the fundamental task for the modelling of OB. With the formed guidelines for monitoring and data collection for occupant behavior and occupancy in Annex 66 (Yan et al. 2017), the plan of a global occupant database for different building types and typical spaces has been identified and proposed in the process of Annex 79, providing research fundamental for developing guidelines and recommendations for occupant-centric design (O’Brien et al. 2020).

1.3 Objectives of this study

Summing up the above, the district-scale occupancy modelling poses challenges from the perspectives of sensing, modelling, interpretation and prediction. The research question is that how to take full advantages of data sensing and statistics technics for upscaling occupancy modelling from building level to district scale. To deal with it, this paper proposes the modelling methods of inter-building movements. The methods utilize multiple distribution fittings and Bayesian network to upscale the event description methods from inter-zone movement events (building-level) to inter-building movement events (district-level). This study provides a framework on the application of the proposed modelling methods for a university campus in the suburbs of Shanghai. The application starts with the campus-scale occupancy data measurement. By analyzing the collected campus-scale occupancy data, this study performs the definitions and descriptions of various inter-building movement events. So that, the campus-scale stochastic process
of occupancy could be realized by MC simulation method on the basis of the defined inter-building movement events for different kinds of students. To validate the feasibility of the proposed modelling, this paper conducts a comparative analysis between the simulation results and the measured occupancy data. Furthermore, the impact of the inter-building movement modelling on building energy demand is evaluated by applying the simulated occupancy results to the library in this campus.

In this paper, the research framework and modelling methods are introduced in Section 2. Section 3 illustrates the application of the proposed methods for the campus in Shanghai as a case study, with discussion on comparative results and key findings. Finally, current study is concluded in Section 4.

2 Methods

2.1 Overview of research framework

Simulation for building-level occupancy can provide good methodological basis for this study. The well-known events mechanism at the building level has classified two kinds of inter-zone movements, (1) typical movement event, such as on/off work and lunch rest, and (2) stochastic movement between different rooms (Wang et al. 2011). The geometric distribution and stationary stochastic process are widely used for the mathematical description of these two kinds of inter-zone movements in a building. As for the simulation process, MC simulation and priority rules of multiple events are applied to perform the complete stochastic process that generating massive occupancy states step by step.

This study initiates the extrapolation of occupancy modelling methodology from building level to district scale through proposing the modelling methods of inter-building movements. Figure 1 illustrates the research framework to apply the extrapolated occupancy modelling methods to a university campus in Shanghai. This paper attempts to establish mathematical profiling for dominated and stochastic inter-building transition events for the district-scale occupancy simulation. During the exploration on upscaling simulation methods, the multiple distributions and Bayesian probability...
network are combined to characterize the dominated events of inter-building movement in the campus. Relative to simulation for building-level occupancy, the proposed method focuses on the transition process/probability between one building to another one, rather than that between one room to another one. The modelling for campus-scale concerns about which building the occupant locates in, and it is ignored that which room/area is or will be occupied in a specific building.

By taking advantages of smart sensing and monitoring technology, this study collected the campus-scale occupancy data, including GIS-based real-time occupancy tracking data from student volunteers, and occupancy data obtained from field counting and BAS for several functional campus buildings. With clustering analysis on the campus-scale occupancy data, five patterns of inter-building movement have been defined by the dominated inter-building movement events for each pattern. The varied patterns can facilitate more complete modelling for campus occupancy, by considering the diversity of students' campus life. In the case study, the simulation for campus-scale occupancy also considers the certain stochasticity of inter-building movement by means of the stationary stochastic process, which is similar to the building-level modelling. Based on the combination of quantitative descriptions for dominated events and certain stochasticity of inter-building movement, this paper performs the stochastic simulation for the campus district, using MC models. The simulation results are then validated with the campus-scale occupancy measurement data. Finally, the impact of inter-building movement modelling methods on building energy demand is evaluated for the library building, taking the deterministic occupancy schedules suggested by current building design standard as a baseline.

2.2 Simulation methods for campus occupancy modelling

In the context of campus district, the inter-building movement could be driven by two aspects: (1) diverse spatiotemporal usage demand influenced by individual class or scientific work plan and operational schedules of mixed-used buildings; (2) stochasticity and uncertainty of inter-building movement. Comprehensive effects of the two aspects on each student eventually form the spatiotemporal usage demand with the campus district. Obviously, the former aspect is relatively deterministic and is possible to form certain patterns of inter-building movements.

For building-level occupancy modelling, statistical distributions are the most commonly used to fit the presence and absence period of occupant movement events within a building, e.g. lunch break. If we use a two-state MC to represent the absorption process of a typical enter/exit event, such as the go-for-lunch event in Eq. (1), the $p_{01}$ reflects the probability of go-for-lunch event occurring at the $n^{th}$ timestep in the lunch period.

$$p_{\text{go lunch}} = \begin{bmatrix} p_{00} & p_{01} \\ 0 & 1 \end{bmatrix}$$

(1)

These typical events have their own start and end time. For example, the go-for-lunch event usually has the start time of 11:00 and the end time of 12:30 for a restaurant building. If we assume that $p_{00}$ and $p_{01}$ are time-independent in the lunch period, the probability of go-for-lunch event occurring at the $n^{th}$ timestep since the start time of lunch could fit a Geometric distribution. Except for Geometric distribution, Poisson (Wang et al. 2005; Page et al. 2008; Wang et al. 2011) and Gaussian (Gilani and O’Brien 2018; Ding et al. 2021) are also commonly used statistical distributions. Based on the distribution fitting, the inhomogeneous two-state Markov chain model with transition probabilities of arriving, leaving and staying states could be generated and the model is capable of reproducing the important characteristic of occupancy for campus buildings during the morning arrival time and the evening departure time (Page et al. 2008). The researchers in Wang et al. (2011) had pointed out that such processes also may be utilized to coupling movement within buildings, depending on the building types, the geographic locations, and occupant preferences, which exactly is the key problem that this study attempts to deal with for the inter-building occupancy modelling and application.

In this study, three commonly-used stochastic process methods, Poisson, Gaussian and Power distribution, are applied as the fitting models to the absorption process of enter/exit events for campus buildings. In the application of multiple distributions, the actual occupancy data shows that $p_{01}$ is time-dependent, which is different with the traditional application of Geometric distribution. For example, the dominate movement event of lunch arrival for the restaurant building could be described by a Power distribution, and the 1st course arrival at the lecture building is described by that of a Gaussian distribution. At the timestep that closer to the noon time (12:00) or class start time (8:30), it has higher transition probability ($p_{01}$) of arriving at the restaurant building or lecture building. Besides, this event may has different start/end time or absorption speed for different pattern of students. So that, the distribution fitting of the inter-building movement events will be separately conducted for different pattern of students.

Obviously, the distribution fitting method aims to describe the absorption process of occupants' arrival and
departure events for one of campus buildings. Except for this, it is essential to know about the stochastic directions and transition probabilities that connected with other buildings. If we take the occupants’ arrival event at lunch time for the restaurant building as an example, a complete occupancy modelling not only can simulate the absorption process of occupants’ arrival for lunch, but also can simulate which buildings the occupants came from and how many occupants came from each of them for lunch. Similarly, the simulation for occupants’ departure event after lunch time not only has the anti-absorption process of occupants’ departure from the restaurant building, but also involves which buildings the occupants will transfer to and how many occupants will go to each of them. This study has adopted Bayesian network methods to quantitatively describe the above stochastic directions and transition probabilities. The network is centered by one campus building, such as the restaurant building, and the center building could be regarded as a traffic hub of occupant streams, including occupants coming from or going to other buildings. The current study established the Bayesian networks by conditional transition probabilities, as formulated by Eqs. (2) and (3), using real-time occupancy tracking data.

\[
P(\text{location}_{t,j}|\text{location}_{t-1,i}) = \frac{P(\text{location}_{t,j}) \cdot P(\text{location}_{t-1,i}|\text{location}_{t,j})}{\sum_{i \in S} P(\text{location}_{t,j}) \cdot P(\text{location}_{t-1,i}|\text{location}_{t,j})} \quad (2)
\]

\[
P(\text{location}_{t-1,i}) = \frac{P(\text{location}_{t,j}) \cdot P(\text{location}_{t-1,i}|\text{location}_{t,j})}{\sum_{j \in S} P(\text{location}_{t,j}) \cdot P(\text{location}_{t-1,i}|\text{location}_{t,j})} \quad (3)
\]

To describe how to apply the proposed modelling methods in the stochastic process simulation for a campus, Figure 2 shows the simulation flowchart. The simulation needs key parameters of inter-building movement events as the inputs. Dominated events, such as arriving for or leaving from restaurant, lecture or lab, etc., are considered in this study. Each event has the characterized parameters including start time, end time, peak time, student proportion of each pattern, as well as the event-related transition probability matrixes.

If there has 7 main buildings/areas in the campus, we will establish a 7-D matrix \((\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7)\), and the value of \(\pi_i\) represents the number of students who are located at the \(i^{th}\) building/area. The simulation starts with primary location setting of \(N\) students \((N, 0, 0, 0, 0, 0, 0)\) at \(t_0\). It means that all of \(N\) students are located at the 1st building/area at the beginning of the simulation. We set the time interval of 15 min for a step in the simulation. Thus, for the next step, at \(t_1\), the location statement of each student will be updated based on the random MC simulation. The MC simulation utilizes the movement events and transition matrixes for the next step \(t_1\) to perform the random process generation. For each step, the proportions of occupants for different patterns keep the same as the clustering results. After each step, the location situation \((N, 0, 0, 0, 0, 0, 0)\) will be updated to \((n_1, n_2, n_3, n_4, n_5, n_6, n_7)\), which means the occupancy number at the \(i^{th}\) building/area changes to \(n_i\). The step-by-step simulation will repeat until all the defined events in a day are completed. In this way, each time of the completed simulation can achieve occupancy schedules of at least 15 min interval for each of the 7 main buildings/areas.

3 Case study

3.1 Data acquisition of campus-scale occupancy

The current case study has been implemented to a university
campus located at the suburb of Shanghai. The university campus can be divided into administrative service area, public teaching area, academic lab area, sport activity area, dormitory area, catering area, and public green area. This campus is dominated by students, including 6,154 undergraduate and 6,011 graduate students in total. Here, we focus on the presence and movement of the students. Majority of students have strong time-dependent study-life styles due to individual course schedules or scientific work plans. Different from central business district with commercial buildings, the campus district is a relatively close-off area with fully functional buildings/areas serving for fruitful study-life activities.

By the way of installing an APP called Ovital interactive map on the cellphones of volunteer students, this study has obtained the GIS-based real-time occupancy tracking data from 193 volunteer students during two weeks in September 2020. Figure 3 shows the screenshot of Ovital interactive map. It shows the partial map of the case campus and the blue lines with direction of arrows represent the movement of this student within a day in the campus tracked by the APP. The volunteer students are required to keep the phone with himself/herself all the day during the test period, and open the Ovital app every morning when they get up so that it runs in the background for the whole day until they go sleep and close it. Ovital records the location data at every moment that the student moves. The raw data of real-time location includes the longitude, latitude and altitude, along with the timestamp. The accuracy of GIS location data has the distance error of less than 5 m. The raw location data of each volunteer has been stored in a file of .ovobj that can be read with Python. An algorithm called PN POLY (point-in-polygon algorithm) is applied to the raw location data to estimate which building or area the student locates, by comparing the relative location between a polygon of the building or area with the student’s location point. For the convenience of following data analysis, the raw data is processed by interpolation with an interval of 5 min, and each data record adds the label of the located building or area within the campus.

The representativeness of the tracking data to the entire campus is an important concern for the current study on campus-scale occupancy modelling. To address this issue, we’ve conducted a campus-wide volunteer recruitment by releasing advertisements for paid volunteers through the university public social media. From the perspective of statistics, the sampling ratio for the whole campus having around 12,165 students should be no less than 1% to ensure the samples’ representativeness for the whole campus population. It means the number of samples should be more than 122 students (12,165 × 0.01 ≈ 122). Our study successfully recruited 193 volunteer students, which is statistically enough. Besides, we checked the proportion structure of their student identity. It has been found that the proportions of undergraduate, master and PhD students in the 193 volunteer students are 50%, 35%, and 15%, respectively, which are similar to those in the whole campus population. Also, they come from about 10 academic colleges and their majors involve mechanical engineering, transportation engineering, material engineering, electronic

Fig. 3 User interface of GIS-based tracking data monitor APP, Ovital interactive map
Therefore, the tracking data from the volunteers could reflect great diversity of the campus life.

At the same time, this study obtains the second dataset of occupant flows of several typical buildings by two means, building automation system (BAS) of library and colleges, and counting of occupants in the fields such as building entrance and exit. Table 2 provides the details of the data source, including data measurement method and period, monitoring time and the building types. This study carried out the field counting of occupants for four dormitory buildings, two lecture halls, and one restaurant building for one week in September 2020. Each of the seven buildings has no more than two entrances or two exits for the healthy management during Covid-19 period. The timings of field counting depend on the building usage situation, such as lunch and dinner time for restaurant building, time of before and after classes for lecture halls, almost whole day for dormitory buildings. During these time periods, each research member presses the count button on a counter (shown in Table 2) when every student pass by and the counting number increases by one and the accumulated counting number is recorded by every 5 min. Each entrance/exit has two research members for double check on the manual counting number to minimize error. Through the monitoring of occupant flows at the above key timepoints for inter-building movements, the dataset can cover the main movements of students in this case campus.

Besides, this study has obtained data from the BAS that stores the in/out records of campus card for the library and a college building. All the students in the campus have their own campus cards that identify their student ID number, and the card is used to purchase in the restaurant, open the entrances having access control system to the apartment, library and college building. The access control system can be opened by one-time swipe of a card and only allow one people to pass by at one time. The card can be regarded as a necessity and needs to be taken with a student to almost everywhere in this campus. Understandably, some situations, like that multiple people get in with only one card, or someone borrows the card owned by another student, are rare and not allowed within the campus. The above card system is widely used in many university campuses of China. These card records are reliable data source in the study on the campus occupant behavior research. In our study, the raw records of the library include the timestamp of the card swipe, the affiliated college of the identified student ID number, the label of get in or go out. And the raw records of the college lab building include the number of people who get in and go out of the building during each of the past hour.

### 3.2 Patterns of dominated inter-building movement

GIS-based real-time occupancy tracking data includes 2702 daily location trackings (193 × 14 = 2702) collected from 193 students for 14 days. Each tracking includes longitude and latitude of the location, as well the label of which building/area the student locates at, for every 5 min in a day. For the convenience of pattern analysis, each tracking is processed to a sample of 16-D occupancy-related features. Figure 4 shows an example of 16-D feature extraction for a student on a weekday. The residence time ratio of building $i$ is defined as the proportion of the time duration that a student stays at building $i$ in a day, as formulated by Eq. (4), in which the $TL_i$ refers to the length of time that the test student stays in the building $i$, and $T_E$, $T_S$ are the time of the last and the first movement monitored in the test day, respectively.

$$\text{Residence time ratio}(i) = \frac{TL_i}{T_E - T_S}$$

| The use of building(s)                                      | Monitoring time                                                                 | Measurement period        | Data collection method                  |
|-------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------|----------------------------------------|
| 4 dormitory buildings for undergraduate and graduate students | 6:00–24:00                                                                      | 1 week                   | Field counting of occupants at building entrance and exits by counters |
| Restaurant building with cafeteria                          | Lunch time (10:30–13:00), Dinner time (16:30–19:00)                             | (Sep 24–Sep 28, 2020)    |                                        |
| 2 teaching and lecture halls                                 | Before the first class (7:00–8:10), Intervals between classes (9:30–10:10; 11:30–13:40; 15:00–15:40; 17:00–19:10), After the last class (20:30–21:30) | 1 month                   | BAS with ID card system records       |
| 1 college lab building                                      | 0:00–24:00                                                                      | (Sep 20–Oct 20, 2020)    |                                        |
| Library building                                            | 0:00–24:00                                                                      | 1 month                   | BAS with ID card system records       |
The residence time ratio of a specific building is a considerable character due to the fact that the more time an occupant stays in a building will result in relatively more end use and energy consumption. Except for that, other considerable characters include arrival frequency of each building and the timestamps that the first (\(T_S\)) and last (\(T_E\)) movement occur.

With the sample size of 2702 16-D dataset of inter-building movements in the campus, this study has recognized five kinds of campus study-life patterns by \(K\)-means clustering (Wang et al. 2013; Zhang and Wen 2019). At the beginning of clustering analysis, each of the 16-D features is normalized to a range of 0 to 1, as formulated by Eq. (5). The normalization can eliminate the effect of varying scales and units among the 16-D features on the clustering results, and each of the normalized feature can apply the same weight in the distance metric during the clustering process. For the \(j\)th feature (1 \(\leq j \leq 16\)):

\[
X_{\text{norm},i} = \frac{X_{\text{raw},i} - \min(X_{\text{raw}})}{\max(X_{\text{raw}}) - \min(X_{\text{raw}})} \quad (1 \leq i \leq 2702)
\]  

Each column of the numbers in Table 3 means the normalized 16-D center for each of the five resulted clusters. Based on the normalization rule, we can understand in this way, the larger residence time ratio of the building means that this pattern of students spends more time in the building during the day time, than other patterns. If we take the first feature in Table 3, residence time ratio of Dormitory, as an example, students of Pattern 1 spend the most residence time in the Dormitory and those of Pattern 5 spend the least time in the Dormitory. And the increased order of the residence time in the Dormitory starts from Pattern 5 (value = 0), followed by Pattern 4 (value = 0.40), Pattern 2 (value = 0.50), Pattern 3 (value = 0.63), then Pattern 1 (value = 1).

It is necessary to figure out that which kinds of students constitute each pattern. Among the monitored 193 samples, the student proportions of five patterns are 17%, 20%, 19%, 24% and 20%, respectively. As shown in Figure 5, the majorities of Patterns 1, 2, and 3 are undergraduates, that of Patterns 4 and 5 are graduates. The proportion structure demonstrates that the diversity of inter-building movement is considerable on the campus scale and identical schedules suggested by current domestic building design standards need to be improved and expanded. To extract dominated events for each resulted pattern, we utilized the GIS-based tracking logs to count the inter-building movement frequencies between every two buildings during the test period. Figure 6 shows partial frequencies of inter-building movement within the campus for students of Patterns 2 and 4, and the red arrows and numbers refer to the inter-building movement that having an average frequency per student of more than 5. As an example, the student count of Pattern 2 is 39, and the frequencies more than 195 (39 \(\times 5\)) are red highlighted in Figure 6(a). Obviously, students of Pattern 2 are centered with Lecture hall, Dormitory, and Restaurant buildings, while students of Pattern 4 are centered with Lab, Dormitory, and Restaurant buildings in their campus life. The former mainly stands for a study-life style of undergraduates, the latter is for graduates in this campus. Here, the buildings that students frequently enter and exit are defined as the movement centers, and the dominated movement events are defined as the inter-building movement between two of them. As listed in Table 4, this study provides...
the definition on the five patterns of inter-building movements, including the portrait of majority of students, and the possible inter-building movements between the movement centers as the dominated events in their regular campus life, for each pattern.

To provide a better understanding of different campus life, Figure 7 illustrates hourly spatiotemporal distribution of the sampled students on a workday, separately for graduates and undergraduates. Undergraduates spend great proportion of daytime at the lecture hall for their GPA classes, see the light blue part in Figure 7(b). While, graduates have much less classes and they spend more time on their research work at the college lab, see the orange part in Figure 7(a). For both undergraduates and graduates, the peak time of lunch/dinner at the restaurant (green part in Figure 7) perfectly matches with the valley time of class or work breaks. Besides, graduates take more time outside campus than undergraduates, which is reasonable since graduates may join internship for career and need to commute outside campus.

Though there exist different patterns for campus life of the sampled students, the movement centers that those students frequently enter and exit basically consist of the dormitory, the restaurant, the lab, the lecture hall and the library, which are public and open to all the students. This study utilizes the data collected from field counting and BAS to realize the occupancy trends during a day for the movement centers.

As mentioned in Section 3.1, the field counting method records the counting numbers of student entered and exited the building for every 5 min. As formulated by Eq. (6), $I_t$ and $I_{t-1}$ represent the accumulated numbers of occupants who entered the building at $t$ and $t-1$ time, respectively. $O_t$ and $O_{t-1}$ represent the accumulated numbers of students who exited the building at $t$ and $t-1$ time, respectively. Therefore, $N_t$ and $N_{t-1}$ could represent the occupancy number of students in the building at $t$ and $t-1$ time, respectively.

| Features                                | Pattern 1 | Pattern 2 | Pattern 3 | Pattern 4 | Pattern 5 |
|-----------------------------------------|-----------|-----------|-----------|-----------|-----------|
| Residence time ratio of L1. Dormitory   | 1.00      | 0.50      | 0.63      | 0.40      | 0.00      |
| Residence time ratio of L2. Restaurant  | 0.00      | 0.37      | 0.72      | 1.00      | 0.62      |
| Residence time ratio of L3. Lab         | 0.14      | 0.00      | 0.03      | 1.00      | 0.26      |
| Residence time ratio of L4. Library     | 0.17      | 0.23      | 0.37      | 0.00      | 1.00      |
| Residence time ratio of L5. Gym         | 0.00      | 0.45      | 1.00      | 0.47      | 0.01      |
| Residence time ratio of L6. Lecture hall| 0.09      | 1.00      | 0.73      | 0.00      | 0.08      |
| Residence time ratio of L7. Outside campus| 0.00  | 0.02      | 0.38      | 0.28      | 1.00      |
| Arrival times of L1. Dormitory           | 0.00      | 1.00      | 0.36      | 0.38      | 0.69      |
| Arrival times of L2. Restaurant          | 0.45      | 0.92      | 1.00      | 0.68      | 0.00      |
| Arrival times of L3. Lab                 | 0.02      | 0.00      | 0.15      | 1.00      | 0.20      |
| Arrival times of L4. Library             | 0.16      | 0.22      | 0.27      | 0.00      | 1.00      |
| Arrival times of L5. Gym                 | 0.00      | 0.23      | 1.00      | 0.26      | 0.08      |
| Arrival times of L6. Lecture hall        | 0.17      | 1.00      | 0.63      | 0.00      | 0.11      |
| Arrival times of L7. Outside campus      | 0.00      | 0.05      | 0.67      | 0.15      | 1.00      |
| Time of the first movement in a day      | 1.00      | 0.00      | 0.36      | 0.40      | 0.48      |
| Time of the last movement in a day       | 0.00      | 0.59      | 0.70      | 0.84      | 1.00      |

Fig. 5 Proportions of undergraduates and graduates for five patterns.
Figure 8 displays the example of the occupancy changing trends of $N_t$ for a dormitory building, and those for the restaurant during lunch and dinner time on Sep 24 in 2020 (Thursday). It shows that the dormitory buildings have occupancy schedules of U-shape. The valley-to-peak ratio basically stays at the range of 26%–29%. Besides, the restaurant building has obvious double-peak profile. The ratio of dinner-period occupancy to lunch-period occupancy is about 0.7–0.8 on the weekday, while that is about 0.5 at the weekend.

Based on the BAS card records, Figure 9 also provides the occupancy changing trends of $N_t$ for the lab and the library. The occupancy schedules all have three obvious peaks that separately occur at about 10:30 am, 15:30 and 19:30, and two

$$N_t = N_{t, in} - N_{t, out}$$

where $N_{t, in} = I_t - I_{t-1}$ and $N_{t, out} = O_t - O_{t-1}$

(a) Students of Pattern 2

(b) Students of Pattern 4

Fig. 6 Partial frequencies of inter-building movement for students of different patterns

Table 4 Student portrait and dominated events of five patterns

| Patterns            | Portrait                        | Movement centers                  | Dominated inter-building movement events                                      |
|---------------------|--------------------------------|-----------------------------------|--------------------------------------------------------------------------------|
| Pattern 1           | Students staying in dormitory most of time | Dormitory and Restaurant          | Breakfast arrival, Breakfast departure, Lunch arrival, Lunch departure, Dinner arrival, Dinner departure, Back to dormitory |
| Pattern 2           | Students having lots of courses | Dormitory, Restaurant and Lecture hall | Breakfast arrival, Breakfast departure, Lunch arrival, Lunch departure, Dinner arrival, Dinner departure, the $N^{th}$ course arrival, the $N^{th}$ course departure, Back to dormitory |
| Pattern 3           | Students having less courses   | Dormitory, Restaurant and Lecture hall | Breakfast arrival, Breakfast departure, Lunch arrival, Lunch departure, Dinner arrival, Dinner departure, the $N^{th}$ course arrival, the $N^{th}$ course departure, Back to dormitory |
| Pattern 4           | Students staying in lab most of time | Dormitory, Restaurant and College lab | Breakfast arrival, Breakfast departure, Lunch arrival, Lunch departure, Dinner arrival, Dinner departure, the $N^{th}$ arrival at lab, the $N^{th}$ departure from lab, Back to dormitory |
| Pattern 5           | Students often out-campus (for intern) | Dormitory, Restaurant, College lab, and outside campus | Breakfast arrival, Breakfast departure, Lunch arrival, Lunch departure, Dinner arrival, Dinner departure, the $N^{th}$ arrival at lab, the $N^{th}$ departure from lab, Outside campus, Back to campus, Back to dormitory |
valleys at about 11:30 and 17:30. The time of valleys just match the two peaks of lunch and dinner periods. These practical occupancy trends are different from the commonly-used office schedules suggested by some building design standards, which usually have the two-phase shape and underestimate the building usage at night. The results indicate that the hourly occupancy trends for both weekdays and non-workdays have similar peak and valley periods. Different from common office buildings, the occupancy trends of campus buildings are more stable and don’t have obvious difference between weekdays and weekends, especially for graduate students who have few classes to take on weekdays. Furthermore, Figure 10 provides the dynamic trends of the average $N_t$ for each day. The dynamic changes from day to day for the lab and the library are similar. Both of them increase gradually in the public holidays (National day of China), followed by relatively stable trends for the next weekdays. It’s reasonable that during the public holiday period, the day closer to the first workday after the vacation has more students occupied in the lab and library. After the holiday period, each of the weekdays has the similar average number of occupants.

### 3.3 Campus-scale Bayesian network of inter-building movement

As a result, Table 5 lists the optional dominated events and the fitted stochastic process distributions. For the dominated events, such as breakfast/lunch/dinner, classes, lab-centered and dormitory-centered movements, this study applies three distributions (Power law, Poisson, and Gaussian) to fit the arrival and departure processes that centered on the dominated functional buildings for different study-life patterns. Different distribution method is used to describe the temporal statistics of arrival or departure occupants at each timestep in the specific event period. Human dynamics area has a traditional assumption that its temporal statistics are uniform and stationary, which can be properly described by a Poisson process. Accordingly, the inter-event time distribution should have an exponential tail, and the Poisson distribution has been widely used to characterize the statistical regularity of occurrence frequency of specific event in certain timesteps, e.g. the modelling of traffic flow patterns, accident or estimation of congestion caused blocked calls (Barabási 2005). In the Poisson method, the probability/frequency of the specific event is almost uniform for each timestep and the inter-event time interval is similar, which indeed matches
Poisson distribution method still is the most commonly used occupancy modelling method with the event mechanism at the building level, due to its representation of uniform random probability. Except for Poisson distribution, an increasing number of recent scholars indicate that the timing of human actions systematically deviates from the Poisson prediction, the inter-event times being better approximated by a heavy-tailed distribution, such as a power law distribution. The power law distribution has more capability of characterizing the bursts of rapidly occurring events separated by long periods of inactivity. The researchers pointed out that the bursty nature of human behavior is a consequence of a decision-based priority and this kind of tasks would be executed rapidly. Relatively, the priority blind execution could be well approximated by uniform inter-event statistics (Barabási 2005). Based on the time-dependent occupancy measurement data profiles, this study takes advantages of key findings in human dynamics area and adopts power law and Gaussian distributions to the modelling of arrival and departure movements centered on lunch/dinner, classes and lab work, those have decision-based priority for a student’s campus life and show obvious occupant bursts and aggregations from the event-related measurement data.

Here we take the students of Pattern 2 (undergraduate students having lots of courses) as an example and provide the developed Bayesian probability network for them, as shown in Figure 11. This network covers their dominated events of daily study-life portrait and the corresponding stochastic probability of inter-building movements. We built a typical timeline that made up of main inter-building movement events orderly having the transfer centers of restaurant and lecture buildings for the students of Pattern 2. In the beginning of a workday, the most possible inter-building movement events are to have breakfast ($t_b$) in the restaurant and to have the first class ($t_{k_1}$) in the lecture building. For the restaurant, all of students come from dormitory with the condition that close-off management keeps all students living inside the campus. After the breakfast, the proportions of students who are going to dormitory, lab, library and lecture buildings are 31%, 15%, 12% and 42%, respectively, see the probability network at $t_b'$ period.

Similarly, when the Lecture building is regarded as the center of inter-building movement events, the proportions of students that come from other buildings to take GPA classes and that transfer to other buildings after their classes in this network make this kind of students’ campus life interpretable and quantitative for campus-scale occupancy modelling.

Fig. 9 Occupancy profiles of college lab and library

Fig. 10 Dynamic trends of daily mean occupancy for college lab and library

with the time-dependent occupancy measurement data of the breakfast arrival and departure events in this study. Poisson distribution method still is the most commonly used occupancy modelling method with the event mechanism at the building level, due to its representation of uniform random probability. Except for Poisson distribution, an increasing number of recent scholars indicate that the timing of human actions systematically deviates from the Poisson prediction, the inter-event times being better approximated by a heavy-tailed distribution, such as a power law distribution. The power law distribution has more capability of characterizing the bursts of rapidly occurring events separated by long periods of inactivity. The researchers pointed out that the bursty nature of human behavior is a consequence of a decision-based priority and this kind of tasks would be executed rapidly. Relatively, the priority blind execution could be well approximated by uniform inter-event statistics (Barabási 2005). Based on the time-dependent occupancy measurement data profiles, this study takes advantages of key findings in human dynamics area and adopts power law and Gaussian distributions to the modelling of arrival and departure movements centered on lunch/dinner, classes and lab work, those have decision-based priority for a student’s campus life and show obvious occupant bursts and aggregations from the event-related measurement data.

Here we take the students of Pattern 2 (undergraduate students having lots of courses) as an example and provide the developed Bayesian probability network for them, as shown in Figure 11. This network covers their dominated events of daily study-life portrait and the corresponding stochastic probability of inter-building movements. We built a typical timeline that made up of main inter-building movement events orderly having the transfer centers of restaurant and lecture buildings for the students of Pattern 2. In the beginning of a workday, the most possible inter-building movement events are to have breakfast ($t_b$) in the restaurant and to have the first class ($t_{k_1}$) in the lecture building. For the restaurant, all of students come from dormitory with the condition that close-off management keeps all students living inside the campus. After the breakfast, the proportions of students who are going to dormitory, lab, library and lecture buildings are 31%, 15%, 12% and 42%, respectively, see the probability network at $t_b'$ period.

Similarly, when the Lecture building is regarded as the center of inter-building movement events, the proportions of students that come from other buildings to take GPA classes and that transfer to other buildings after their classes in this network make this kind of students’ campus life interpretable and quantitative for campus-scale occupancy modelling.
### Table 5 Distribution fitting of dominated inter-building movement events

| Movement center | Dominated event options                      | Fitted distribution       | Event period | Occupancy measurement data |
|-----------------|---------------------------------------------|---------------------------|--------------|----------------------------|
| Restaurant      | Breakfast arrival, breakfast departure       | Poisson distribution      | $t_b, t_b'$  | Time-dependent number of occupants arrived for breakfast at each step in $t_b$ period |
| Restaurant      | Lunch arrival, lunch departure               | Power distribution, Poisson distribution | $t_l, t_l'$ | Time-dependent number of occupants arrived for lunch at each step in $t_l$ period |
| Restaurant      | Dinner arrival, dinner departure             | Power distribution, Poisson distribution | $t_d, t_d'$ | Time-dependent number of occupants arrived for dinner at each step in $t_d$ period |
| Lecture hall    | The $N^{th}$ course arrival, the $N^{th}$ course departure | Gaussian distribution, Poisson distribution | $t_c, t_c$ | Time-dependent number of occupants arrived for class at each step in $t_c$ period |
|                 |                                             |                           |              | Time-dependent number of occupants left from class at each step in $t_c$ period |
3.4 Validation and application of campus-scale inter-building occupancy models

This case study has utilized key event parameters and performed the campus-scale stochastic process with MC simulation, according to the simulation flowchart shown in Figure 2. To eliminate the uncertain impact of single random simulation, 100 times of simulation have been performed and averaged for each time step. Figure 12 shows the averaged results of 100 times simulation for campus buildings, such as Restaurant, Dormitory and Library (orange lines). To validate the model feasibility, Figure 12 also compares the simulated results of occupancy schedules with the occupancy measurement data (blue lines). It has demonstrated that the simulated occupancy profiles basically align with practical situation.

For the restaurant in Figure 12(a), occupancy profiles for measurement and simulation have consistent peak time, around 12:00, during the lunch period (see the red dash line). After that, the simulated occupancy decreases little slower than the measurement. For the dormitory in Figure 12(b), the occupancy curves of both simulation and measurement have the U-shape with two slight valleys that separately around lunch and dinner time (see the red dash lines). As we can see in the Bayesian probability networks (Figure 11), among the several buildings that students come from for lunch or dinner, the proportions of students come from dormitory are up to 31% for lunch and 45% for dinner. At the same time, among the several buildings that students will head to after lunch or dinner, the proportions of students will go to dormitory increase to 75% and 84%, respectively. Besides, the dormitory occupancy curves also have slight peaks between 13:00 to 13:30, which is also supported by the fact that the classes at afternoon begin at 13:30 as the 5th–6th classes during weekdays, and about 69% of students those take the 5th–6th classes come from the

### Table 5 Distribution fitting of dominated inter-building movement events

| Movement center | Dominated event options | Fitted distribution | Event period | Occupancy measurement data |
|-----------------|-------------------------|--------------------|-------------|---------------------------|
| College lab     | The \( N^{th} \) arrival at lab, the \( N^{th} \) departure from lab | 3 phases with \( t_x, t_x', t_c \) | Gaussian distribution | Time-dependent number of occupants arrived for lab at each step in \( t_x \) period |
| Outside campus | Outside campus, back to campus | Poisson distribution | \( t_x, t_c \) | — |
| Dormitory      | Back to dormitory | Poisson distribution | \( t_c \) | — |
Fig. 11 Bayesian probability network for students of Pattern 2 dormitory (Figure 11). The similar situations occur at the times before the 1st–2nd classes which start at 8:00 and after the final classes which ends at 22:00 (see the grey dash lines). The above implies that main inter-building movement events influencing dormitory occupancy profiles are going to the first classes in the morning and in the afternoon, back to dormitory after the final classes in the evening, lunch and dinner as well. Generally speaking, the simulation for inter-building movement models has the ability to reflect the distribution feature and trends of campus-scale occupancy in this study.

Previous energy simulation of campus buildings often takes office occupancy in the simulation of library, lecture hall, and lab due to the lack of directly matched schedules in the design standard (such as GB 50189-2015 Design Standard for Energy Efficiency of Public Buildings in China). The latest version of GB 50189 (2015) provides occupancy schedules for only one kind of campus buildings, lecture hall, and the suggested schedules for lecture hall has the same hourly occupancy ratios with those suggested for office buildings.

From the perspective of application for building energy simulation, this study applies the simulated occupancy schedules in Figure 12(c) as the inputs for the building energy model of the library. At the same time, the hourly energy demand results using the suggested occupancy schedules in GB 50189-2015 are defined as the baseline. Figure 13 provides the comparison between the simulated occupancy schedule with the GB 50189 (2015) occupancy schedule for the library. Figure 14 compares the two simulation
results of hourly energy demand, separately for summer and winter design day. The yellow lines represent the relative difference of hourly simulated energy demand between that using the simulated occupancy schedule in this paper (blue bars) and that using GB 50189-2015 occupancy schedule (green bars, baseline). The results show that, relative to GB 50189-2015 occupancy schedule, the impact of the campus-scale occupancy modelling on the hourly energy demand is considerable. The absolute difference ratios range from −15% to 20%, for the summer and winter design day. The GB 50189-2015 occupancy schedule makes the building energy demand overestimated at daytime (from 9:00 to 18:00), while underestimated at night (after 19:00). Obviously, the simulated occupancy schedule has decreased the peak and sum of simulated occupancy and hourly building energy demand. This implies that more detailed modelling of campus-scale occupancy may reduce the peak load of campus building, which is a key factor for campus energy system design. Besides, the enclosed campus makes students convenient to study or work at night in the library, especially for a campus located at suburb. That’s the possible reason for the higher occupancy ratio at night time, see the blue line in Figure 13. Therefore, it is reasonable that at night time, the simulated energy demand using simulated occupancy schedule is much more than that using GB 50189-2015 occupancy schedule.

This study has demonstrated that the proposed inter-building movement modelling basically match to real occupancy measurement. The significantly relative difference in Figure 14 implies the discrepancy of the current building design with the reality. The possible reason may be the shortage of sufficient survey or measurement data for the subdivided types of building. Good understanding and quantification of students’ inter-building movements can support the further breakdown of building energy usage patterns for different types of campus buildings. For this gap, this study also can provide certain reference for more optional occupancy schedules for more subdivided types of campus buildings.

![Fig. 13 Comparison of simulated occupancy with GB 50189 (2015) occupancy for library building](image)

4 Conclusion

District-scale occupancy modelling has been facing the challenges from the scarcity of district-scale data, modelling methods, as well as simulation application. This study has found an alternative way to explore the methodological upscaling of occupancy modelling from building level to district level. The methodological innovation is that the campus-scale occupancy modelling methods utilize multiple distribution fittings and Bayesian network to upscale the event description methods from inter-zone movement events (building-level) to inter-building movement events (district-level). Relative to simulation for building-level occupancy, the proposed campus-scale methods concern about which building the occupant locates in, and it focuses on the transition process/probability between one building to another one, rather than that between one room to another one in a specific building.

We’ve performed a successful case study on the campus-scale modelling methods of inter-building movements using a university campus in Shanghai. By taking advantages of smart sensing and monitoring technology, this study collected the campus-scale occupancy data, including GIS-based real-time occupancy tracking data from student volunteers, and occupancy data obtained from field counting and BAS for several functional campus buildings. Based on clustering analysis of the campus-scale occupancy data, our work found it become obvious that diversity of students’
campus life when students having different identity live in a close-off campus. This study has recognized five patterns of inter-building movement, and each of them represents the dominated inter-building movement events for one kind of students in their daily campus life. For example, the events of undergraduates are mainly centered with lecture hall, dormitory, and restaurant buildings, while that of graduates are centered with lab, dormitory, and restaurant buildings. The campus-scale simulation results have been confirmed that the inter-building movement modelling can basically reflect the realistic spatiotemporal distribution of campus-scale occupancy. Meanwhile, the relative difference of building energy demand caused by the simulated occupancy schedules implies that, more detailed modelling of campus-scale occupancy can get less energy demand results, relative to using the occupancy schedule suggested by building design standards.

To sum up, this study has led us to conclude that the inter-building movement modelling methods could be an alternative way to simulate campus-scale occupancy with sufficient data. Our research clearly has some limitations. Due to limited funds, this work just conducted the data measurement for around several weeks or one month, and just collected GIS-based real-time tracking data from hundreds of students. Given the relatively short-term occupancy data acquisition in the current case study, we just simulated occupancy schedules for one day to validate the feasibility of the proposed methods. While with this paper as a solid basis, our future research can collect more data to support the occupancy schedule generation for longer time periods. Nevertheless, we believe our work could be a good starting point for the methodological and application upscaling of occupancy modelling from single building level to multiple scales.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (No. 51978481).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Author Contribution

Yiqun Pan, Mingya Zhu, and Zhizhong Huang contributed to the study conception and design. Material preparation, data collection and analysis were performed by Zejun Wu and Mingya Zhu. The first and second draft of the manuscript were written by Mingya Zhu. Previous versions of the manuscript were reviewed and commented by Yiqun Pan and Risto Kosonen. All authors read and approved the final manuscript.

References

Aerts D, Minnen J, Glorieux I, et al. (2014). A method for the identification and modelling of realistic domestic occupancy sequences for building energy demand simulations and peer comparison. Building and Environment, 75: 67–78.

An J, Yan D, Hong T, et al. (2017). A novel stochastic modeling method to simulate cooling loads in residential districts. Applied Energy, 206: 134–149.

Andersen PD, Iversen A, Madsen H, et al. (2014). Dynamic modeling of presence of occupants using inhomogeneous Markov chains. Energy and Buildings, 69: 213–223.

Bætens R, Saelens D (2016). Modelling uncertainty in district energy simulations by stochastic residential occupant behaviour. Journal of Building Performance Simulation, 9: 431–447.

Barabási AL (2005). The origin of bursts and heavy tails in human dynamics. Nature, 435: 207–211.

Benezeth Y, Laurent H, Emile B, et al. (2011). Towards a sensor for detecting human presence and characterizing activity. Energy and Buildings, 43: 305–314.

Bøgger M, Wittchen KB (2018). Estimating the energy-saving potential in national building stocks—A methodology review. Renewable and Sustainable Energy Reviews, 82: 1489–1496.

Carlucci S, de Simone M, Firth SK, et al. (2020). Modeling occupant behavior in buildings. Building and Environment, 174: 106768.

Chen Y, Hong T, Piette MA (2017). Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis. Applied Energy, 205: 323–335.

D’Oca S, Hong T (2015). Occupancy schedules learning process through a data mining framework. Energy and Buildings, 88: 395–408.

Ding Y, Wang Q, Wang Z, et al. (2019). An occupancy-based model for building electricity consumption prediction: A case study of three campus buildings in Tianjin. Energy and Buildings, 202: 109412.

Ding Y, Chen W, Wei S, et al. (2021). An occupancy prediction model for campus buildings based on the diversity of occupancy patterns. Sustainable Cities and Society, 64: 102533.

Ding Y, Han S, Tian Z, et al. (2022). Review on occupancy detection and prediction in building simulation. Building Simulation, 15: 333–356.

Dong B, Andrews B, Lam KP, et al. (2010). An information technology enabled sustainability test-bed (ITEST) for occupancy detection through an environmental sensing network. Energy and Buildings, 42: 1038–1046.

Dong B, Yan D, Li Z, et al. (2018). Modeling occupancy and behavior for better building design and operation—A critical review. Building Simulation, 11: 899–921.

Dong B, Liu Y, Fontenot H, et al. (2021). Occupant behavior modeling methods for resilient building design, operation and policy at urban scale: a review. Applied Energy, 293: 116856.
Du J, Yu C, Pan W (2020). Multiple influencing factors analysis of household energy consumption in high-rise residential buildings: Evidence from Hong Kong. Building Simulation, 13: 753–769.

Fonseca JA, Schlueter A (2015). Integrated model for characterization of spatiotemporal building energy consumption patterns in neighborhoods and City districts. Applied Energy, 142: 247–265.

GB 50189 (2015). GB 50189-2015. Design Standard for Energy Efficiency of Public Buildings in China. Beijing: China Architecture & Building Press. (in Chinese)

Gilani S, O’Brien W (2018). A preliminary study of occupants’ use of manual lighting controls in private offices: a case study. Energy and Buildings, 159: 572–586.

Gu J, Xu P, Pang Z, et al. (2018). Extracting typical occupancy data of different buildings from mobile positioning data. Energy and Buildings, 180: 135–145.

Happle G, Fonseca JA, Schlueter A (2018). A review on occupant behavior in urban building energy models. Energy and Buildings, 174: 276–292.

He M, Lee T, Taylor S, et al. (2015). Coupling a stochastic occupancy model to EnergyPlus to predict hourly thermal demand of a neighbourhood. In: the 14th International IBPSA Building Simulation Conference, Hyderabad, India.

Horni A, Scott DM, Balmer M, et al. (2009). Location choice modeling for shopping and leisure activities with MATSim. Transportation Research Record: Journal of the Transportation Research Board, 2135: 87–95.

Hou H, Pawlak J, Sivakumar A, et al. (2020). An approach for building occupancy modelling considering the urban context. Building and Environment, 183: 107126.

Jin Y, Yan D, Zhang X, et al. (2021). A data-driven model predictive control for lighting system based on historical occupancy in an office building: Methodology development. Building Simulation, 14: 219–235.

Kazas G, Fabrizio E, Perino M (2017). Energy demand profile generation with detailed time resolution at an urban district scale: a reference building approach and case study. Applied Energy, 193: 243–262.

Labeodan T, Zeiler W, Boxem G, et al. (2015). Occupancy measurement in commercial office buildings for demand-driven control applications—A survey and detection system evaluation. Energy and Buildings, 93: 303–314.

Liao C, Lin Y, Barooah P (2012). Agent-based and graphical modelling of building occupancy. Journal of Building Performance Simulation, 5: 5–25.

Mahdavi A, Tahmasebi F (2015). Predicting People’s presence in buildings: An empirically based model performance analysis. Energy and Buildings, 86: 349–355.

Mastrucci A, Pérez-López P, Benetto E, et al. (2017). Global sensitivity analysis as a support for the generation of simplified building stock energy models. Energy and Buildings, 149: 368–383.

Menezes AC, Cripps A, Bouchlaghem D, et al. (2012). Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap. Applied Energy, 97: 355–364.

Mirakhorli A, Dong B (2016). Occupancy behavior based model predictive control for building indoor climate—A critical review. Energy and Buildings, 129: 499–513.

Mohammadi N, Taylor JE (2017). Urban energy flux: Spatiotemporal fluctuations of building energy consumption and human mobility-driven prediction. Applied Energy, 195: 810–818.

Mosteiro-Romero M, Fonseca JA, Schlueter A (2017). Seasonal effects of input parameters in urban-scale building energy simulation. Energy Procedia, 122: 433–438.

Nguyen TA, Aiello M (2013). Energy intelligent buildings based on user activity: A survey. Energy and Buildings, 56: 244–257.

O’Brien W, Wagner A, Schweiker M, et al. (2020). Introducing IEA EBC Annex 79: Key challenges and opportunities in the field of occupant-centric building design and operation. Building and Environment, 178: 106738.

Orehounig K, Mavromatidis G, Evins R, et al. (2014). Towards an energy sustainable community: An energy system analysis for a village in Switzerland. Energy and Buildings, 84: 277–286.

Ouf MM, Park JY, Gunay HB (2021). A simulation-based method to investigate occupant-centric controls. Building Simulation, 14: 1017–1030.

Page J, Robinson D, Morel N, et al. (2008). A generalised stochastic model for the simulation of occupant presence. Energy and Buildings, 40: 83–98.

Park JY, Ouf MM, Gunay B, et al. (2019). A critical review of field implementations of occupant-centric building controls. Building and Environment, 165: 106351.

Qian M, Yan D, An J, et al. (2020). Evaluation of thermal imbalance of ground source heat pump systems in residential buildings in China. Building Simulation, 13: 585–598.

Salimi S, Liu Z, Hammad A (2019). Occupancy prediction model for open-plan offices using real-time location system and inhomogeneous Markov chain. Building and Environment, 152: 1–16.

Stjelja D, Jokisalo J, Kosonen R (2020). From electricity and water consumption data to information on office occupancy: a supervised and unsupervised data mining approach. Applied Sciences, 10: 9089.

Tang R, Wang S, Sun S (2021). Impacts of technology-guided occupant behavior on air-conditioning system control and building energy use. Building Simulation, 14: 209–217.

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

Wang C, Yan D, Jiang Y (2011). A novel approach for building occupancy simulation. Building Simulation, 4: 149–167.

Wang J, Wang S, Chung F, et al. (2013). Fuzzy partition based soft subspace clustering and its applications in high dimensional data. Information Sciences, 246: 133–154.

Wang W, Chen J, Song X (2017). Modeling and predicting occupancy profile in office space with a Wi-Fi probe-based Dynamic Markov Time-Window Inference approach. Building and Environment, 124: 130–142.

Wang J, Sarvi M, Ma J, et al. (2022). A modified universal pedestrian motion model: Revisiting pedestrian simulation with bottlenecks. Building Simulation, 15: 631–644.
Yamaguchi Y, Shimoda Y (2010). District-scale simulation for multi-purpose evaluation of urban energy systems. *Journal of Building Performance Simulation*, 3: 289–305.

Yan D, Hong T, Dong B, et al. (2017). IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings. *Energy and Buildings*, 156: 258–270.

Yang T, Bandyopadhyay A, O'Neill Z, et al. (2022). From occupants to occupants: A review of the occupant information understanding for building HVAC occupant-centric control. *Building Simulation*, 15: 913–932.

Zhang L, Wen J (2019). A systematic feature selection procedure for short-term data-driven building energy forecasting model development. *Energy and Buildings*, 183: 428–442.

Zhang J, Zhou X, Lei S, et al. (2022). Energy and comfort performance of occupant-centric air conditioning strategy in office buildings with personal comfort devices. *Building Simulation*, 15: 899–911.

Zhou X, Liu T, Yan D, et al. (2021). An action-based Markov chain modeling approach for predicting the window operating behavior in office spaces. *Building Simulation*, 14: 301–315.