Simulation-Based Optimization of Energy Consumption and Occupants Comfort in Open-Plan Office Buildings Using Probabilistic Occupancy Prediction Model

Shide Salimi

A Thesis
In the Department
Of
Building, Civil, and Environmental Engineering

Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Civil Engineering) at
Concordia University
Montreal, Quebec, Canada

October 2019

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Shide Salimi

Entitled: **Simulation-Based Optimization of Energy Consumption and Occupants Comfort in Open-Plan Office Buildings Using Probabilistic Occupancy Prediction Model**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Civil Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair
Dr. Abdel Sebak

______________________________ External Examiner
Dr. Qingyan Yan Chen

______________________________ External to Program
Dr. Mohammad Mannan

______________________________ Examiner
Dr. Fariborz Haghighat

______________________________ Examiner
Dr. Fuzhan Nasiri

______________________________ Thesis Supervisor
Dr. Amin Hammad

Approved by ______________________________
Dr. Michelle Nokken, Graduate Program Director

09/12/2019 ______________________________
Dr. Amir Asif, Dean, Gina Cody School of Engineering and Computer Science

ABSTRACT

Simulation-Based Optimization of Energy Consumption and Occupants Comfort in Open-Plan Office Buildings Using Probabilistic Occupancy Prediction Model

Shide Salimi, Ph.D.
Concordia University, 2019

Considering the ever-growing increase in the world energy consumption and the fact that buildings contribute a large portion of the global energy consumption arises a need for detailed investigation towards more effective energy performance of buildings. Thus, monitoring, estimating, and reducing buildings' energy consumption have always been important concerns for researchers and practitioners in the field of energy management. Since more than 80% of energy consumption happens during the operation phase of a building's life cycle, efficient management of building operation is a promising way to reduce energy usage in buildings. Among the parameters influencing the total building energy consumption, building occupants' presence and preferences could have high impacts on the energy usage of a building. To consider the effect of occupancy on building energy performance, different occupancy models, which aim to estimate the space utilization patterns, have been developed by researches. However, providing a comprehensive occupancy model, which could capture all important occupancy features, is still under development. Moreover, researchers investigated the effect of the application of occupancy-centered control strategies on the efficiency of the energy-consuming systems. However, there are still many challenges in this area of research mainly related to collecting, processing, and analyzing the occupancy data and the application of intelligent control strategies. In addition, generally, there is an inverse relationship between the energy consumption of operational systems and the comfort level of occupants using these systems. As a result, finding a balance between these two important concepts is crucial to improve the building operation. The optimal operation of building energy-consuming systems is a complex procedure for decision-makers, especially in terms of minimizing the energy cost and the occupants' discomfort.

On this premise, this research aims to develop a new simulation-based multi-objective optimization model of the energy consumption in open-plan offices based on occupancy dynamic profiles and occupants' preferences and has the following objectives: (1) developing a method for extracting detailed occupancy information with varying time-steps from collected Real-Time Locating System (RTLS) occupancy data. This method captures different resolution levels required for the application of intelligent, occupancy-centered local control strategies of different building systems; (2) developing a new time-dependent inhomogeneous Markov chain occupancy prediction model based on the derived occupancy information, which distinguishes the temporal behavior of different occupants within an open-plan office; (3) improving the performance of the developed occupancy prediction model by determining the near-optimum length of the data collection period, selecting the near-optimum training dataset, and finding the most satisfying temporal resolution level for analyzing the occupancy data; (4) developing local control algorithms for building energy-consuming systems; and (5) integrating the energy simulation model of an open-plan office with an optimization algorithm to optimally control the building energy-consuming systems and to analyze the trade-off between building energy consumption and occupants' comfort. It is found that the occupancy perdition model is able to estimate occupancy patterns of the open-plan office with 92% and 86% accuracy at occupant and zone levels, respectively. Also, the proposed integrated model improves the thermal condition by 50% along with 2% savings in energy consumption by developing intelligent, optimal, and occupancy-centered local control strategies.

ACKNOWLEDGMENT

First and foremost, I wish to express my gratitude to my supervisor Professor Amin Hammad, whose continuous intellectual and personal support, patience and encouragement made this degree possible. His insightful guidance and suggestions were the most valuable help for me, which softened the difficulties of my research study.

I would, also, like to thank the member of my committee, Dr. Qingyan (Yan) Chen, Dr. Fariborz Haghighat, Dr. Fuzhan Nasiri, and Dr. Mohammad Mannan for their valuable inputs and precious time.

This research has been supported by grants from the National Science and Engineering Research Council of Canada (NSERC) and Concordia University Gina Cody School of Engineering and Computer Science. These supports are greatly appreciated.

Furthermore, I would like to use this opportunity to thank Ms. Zheng Liu. I have benefited from her invaluable suggestions and help. Ms. Liu and I collaborated and interacted in the course of my Ph.D. research by conducting multiple lab tests and case studies. Her dedication and support are deeply appreciated.

My sincere thanks go to my close friends, Ms. Roya Kalbassi, Ms. Negar Salimzadeh, Ms. Neshat Bolourian, and Ms. Sara Honarparast who were always there for me with their positivity, which gave me enough energy to walk towards my goals.

I would like to give my special thanks to my beloved parents, Shohreh and Naser, and my dear sister, Dr. Shima Salimi for their endless love and spiritual support and encouragement in all stages of my life. Also, I am thankful to my colleagues for keeping me energized and positive to complete my degree. I will be grateful forever for your love.

Dedicated to my lovely parents, Shohreh and Naser

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
ARHMM	Autoregressive Hidden Markov Model
BBN	Bayesian Belief Network
BECM	Building Energy and Comfort Management
BEMS	Building Energy Management System
BEP	Building Energy Performance
BIM	Building Information Modeling
BLE	Bluetooth Low Energy
CAGR	Compound Annual Growth Rate
CAPM	Context-Aware Power Management
CART	Clustering Algorithm and a Binary Decision Tree
CBM	Cognitive Building Management
CFD	Computational Fluid Dynamics
CPU	Central Processing Unit
DL	Deep Learning
DMTWI	Dynamic Markov Time-Window Inference
DNAS	Drivers-Needs-Actions-Systems
DSOD	Dynamic Spatial Occupancy Distribution
EBC	Buildings and Communities Program
GMM	Gaussian Mixture Model
GP	Genetic Programming
GPS	Global Positioning System
GSM	Global System for Mobile
HMM	Hidden Markov Model
HVAC	Heating, Ventilation, and Air Conditioning
ICT	Information and Communication Technologies
IEA	International Energy Agency
IFC	Industry Foundation Classes
IoT	Internet of Things
KNN	K-Nearest Neighbor

LDR	Light Dependent Resistor
LHMM	Layered Hidden Markov Model
LightWise	LIGHTing evaluation through WIreless Sensor
MCMC	Markov chain-Monte Carlo
ML	Machine Learning
MPC	Model Predictive Control
NMPC	Nonlinear Model Predictive Control
OIM	Occupant Information Modeling
PI	Proportional-Integral
PID	Proportional-Integral-Derivative
PIR	Passive Infrared
POE	Post-Occupancy Evaluation
PS	PresenceSense
RAM	Random Access Measure
RF	Radio Frequency
RFID	Radio Frequency Identification
RSSI	Received Signal Strength Indicator
RTLS	Real Time Locating System
SCOPES	Smart Camera Object Position Estimation System
SMM	Semi Markov Model
SUN	Sensor-Utility-Network
SVM	Support Vector Machine
TD	Time Delay
USSU	User Simulation of Space Utilization
UWB	Ultra-Wideband
WLAN	Wireless Local Area Network
WSN	Wireless Sensor Network
XML	Extensible Markup Language

CHAPTER 1 INTRODUCTION

1.1 Background

It is estimated that world energy consumption will increase by 56% from 2010 through 2040 (SUSRIS, 2013). In 2018, buildings were reported to be responsible for around 40% of the total energy consumption in the United States (21% for residential buildings and 18% for commercial buildings) and assumingly this share will increase around 6% by 2050 based on the AEO2019 Reference case (U.S. EIA, 2019). Similar data are announced by the European Union in terms of energy use in the building sector (26% for residential buildings and 14% for commercial buildings) (EU, 2018). Canada also reported the smaller share for its building sector as 17% for residential buildings and 11% for commercial buildings in 2016 (NRCAN, 2018). Without energy efficiency efforts, this energy use will dramatically increase. Furthermore, more than 80% of energy consumption happens during the operation phase of a building's life cycle (Zhao et al., 2004; Mustapa et al., 2016). Thus, the optimal management of building operations is a promising way to reduce energy usage in buildings (Liang et al., 2016). To provide better building operation and decrease energy consumption, different methods have been studied to make buildings more energy-efficient. This includes implementing intelligent control strategies in building energy-consuming systems, maintaining equipment for achieving maximum efficiency, and increasing the environmental awareness of the occupants. Heating, Ventilation, and Air Conditioning (HVAC) and lighting systems are the prime targets for applying control strategies and energy consumption optimization as these systems account for 71% of the total energy use in commercial and institutional buildings in Canada and almost the same amount in the US (NRCAN, 2016; Wang et al., 2017; Zhu et al., 2017).

The International Energy Agency (IEA), Energy in the Buildings and Communities Program (EBC), Annex 53 recognized the following parameters as the most influential for energy consumption in buildings: (1) climate, (2) building envelope, (3) building energy and service systems, (4) indoor design criteria, (5) building operation and maintenance, and (6) occupant behavior (Annex 53, 2016). Some of these parameters are easy to determine, being related to the physical characteristics of the building (e.g., building size, orientation, construction materials, HVAC system size and type, etc.). On the other hand, some parameters that vary with time are difficult to predict, such as weather and occupancy inputs. The former has been addressed in different research works using reliable data gathered by weather stations and meteorology centers. However, a comprehensive occupancy model is still under development. In addition, it has been proven that since energy-related behavior by occupants has a high impact on all phases of a building's life cycle, consideration of improvements for only the other influential parameters is insufficient to guarantee efficient building energy performance (Pfafferott and Herkel, 2007; Tanimoto et al., 2008). This emphasizes the importance of understanding and considering occupant behavior through proper occupancy modeling.

Occupancy-related information is useful for energy management as well as other areas, such as safety, security, and emergency response. This information includes, but is not limited to, the number, identities, and location of the occupants. According to (Feng et al., 2015), there are four levels of occupancy modeling, which are highly context-dependent. These levels should be determined according to the required granularity of the occupancy models used for different purposes. For instance, a finer level of granularity is needed to apply lighting control strategies, which improve comfort level, since low-resolution occupancy detection can cause occupants' dissatisfaction. Given that HVAC systems need some time to adjust the indoor temperature to a specified target set-point, less accuracy in occupancy detection does not lead to a significant thermal discomfort (Shen et al., 2017). Therefore, four levels of occupancy modeling are defined based on the provided level of detail of occupancy detection as follows: (1) occupancy modeling at the building level considering the number of occupants. This model shows the number of occupants in a building at each time step; (2) occupancy modeling at the space level. This model is defined based on the space state (i.e., occupied or unoccupied) at each time step and is mainly used to control the energy-consuming systems (e.g., lighting) that are not dependent on the number of occupants; (3) occupancy modeling at the space level considering the number of occupants. This model is mainly used to control the HVAC system, which operates based on demand. In this case, the control strategy depends on the number of occupants present in the space at each time step, regardless of who they are; and (4) occupancy modeling at the occupant level. This model tracks each individual; thus, it has the highest level of detail and provides the answers to the following questions: (1) who is in which space? (2) when is that occupant present in the specified space? And (3) what is the occupant doing in the space? Furthermore, post-occupancy evaluation

(POE) is widely used to investigate the effect of occupants' behavior on building performance and energy-saving potentials while maintaining or increasing the occupants' comfort level (Hong and Lin, 2014).

1.2 Problem Definition

As mentioned in Section 1.1, applying intelligent control strategies is one of the methods widely used to make buildings more energy-efficient. Efforts are being made to apply different control strategies improving the buildings' energy consumption while maintaining or increasing the occupants' comfort level. These requirements link the control strategies to the presence of the occupants and their interactions with the building's systems. Hence, evaluating the impact of occupancy-centered control strategies on the buildings' energy-consuming systems is dependent on the occupancy models, which are derived based on the space utilization patterns due to the occupants' behavior. This shows a direct relationship between the space occupancy pattern and the functionality of the building systems. For instance, internal heat gains should be precisely accounted for load calculation and performance analysis of the buildings' HVAC system. Occupancy, use of lights, and other equipment are the main contributions in the internal heat gains, especially in large commercial buildings. However, the uncertainty associated with the internal heat gains that comes from the stochastic nature of the occupancy models leads to overcooling or overheating the space during HVAC systems' operation (Wang et al., 2011). Thus, the optimal control strategies of these systems should be based on the occupancy information in order to meet the needs of the occupants and building energy efficiency simultaneously (Azar and Menassa, 2012). Researchers investigated the effect of the application of different kinds of control strategies on the energy usage of these systems considering the effect of occupancy. However, there are still many challenges in this area of research, mainly related to collecting, processing, and analyzing the occupancy data and the application of intelligent local control strategies, which combine the spatiotemporal variations of the space usage and the occupants' preferences.

Based on the above discussion, the research concerning the energy efficiency of buildings based on the occupant behavior involves the following dimensions as illustrated in Figure 1-1: (1) selecting of the proper type of sensing and monitoring systems including new technologies, such as the Internet of Things (IoT); (2) utilizing a proper occupancy modelling technique to derive deterministic or probabilistic occupancy profiles; (3) applying simulation; (4) using optimization to minimize energy consumption and simultaneously maximize the occupant satisfaction; and (5) applying control strategies to energy-consuming systems.

Monitoring and data collection are the primary steps to develop a detailed occupancy model. However, most of the occupancy detection systems cannot provide the number of occupants and the specific location for each occupant (i.e., the *x* and *y* coordinates of the occupant) when they are used for multi-occupied offices. Therefore, their practicality reduces for open-plan offices that consist of multiple thermal zones (Li et al., 2012). Most of the research works that consider shared multi-occupied offices did not distinguish between different individuals. In addition, all of them lack detailed investigation of the effect of the individual preferences of occupants sharing the same area on the energy consumption of the building. Therefore, there is a need to use proper sensing techniques to distinguish between different occupants in multi-occupied offices and apply their preferences. Post-processing procedures are then used to cleanse the raw data and model the occupancy patterns using different statistical, stochastic, or machine learning (ML) methods. A good occupancy model captures important features pertinent to the occupants and provides a realistic representation of the occupant schedules and behavior (i.e., occupant profiles).

Finally, occupancy models are imported to the energy simulation software to apply control strategies, and ultimately predict building energy consumption (Wang et al., 2011). Energy simulation software requires building parameters and occupancy information as inputs to model the energy performance of buildings. Building energy simulation tools are mature in terms of incorporating building parameters in energy analysis. Some shortcomings are, however, observed regarding occupancy data, which cause large discrepancies in energy usage even between similar buildings with the same characteristics. In order to improve the performance of energy simulation models, the sources of errors regarding the occupancy input data should be investigated. To this aim, the sensitivity of the occupancy prediction models to their input occupancy data needs to be evaluated. The data collection period and the resolution level used to analyze the collected data are two crucial factors for developing accurate occupancy prediction models.

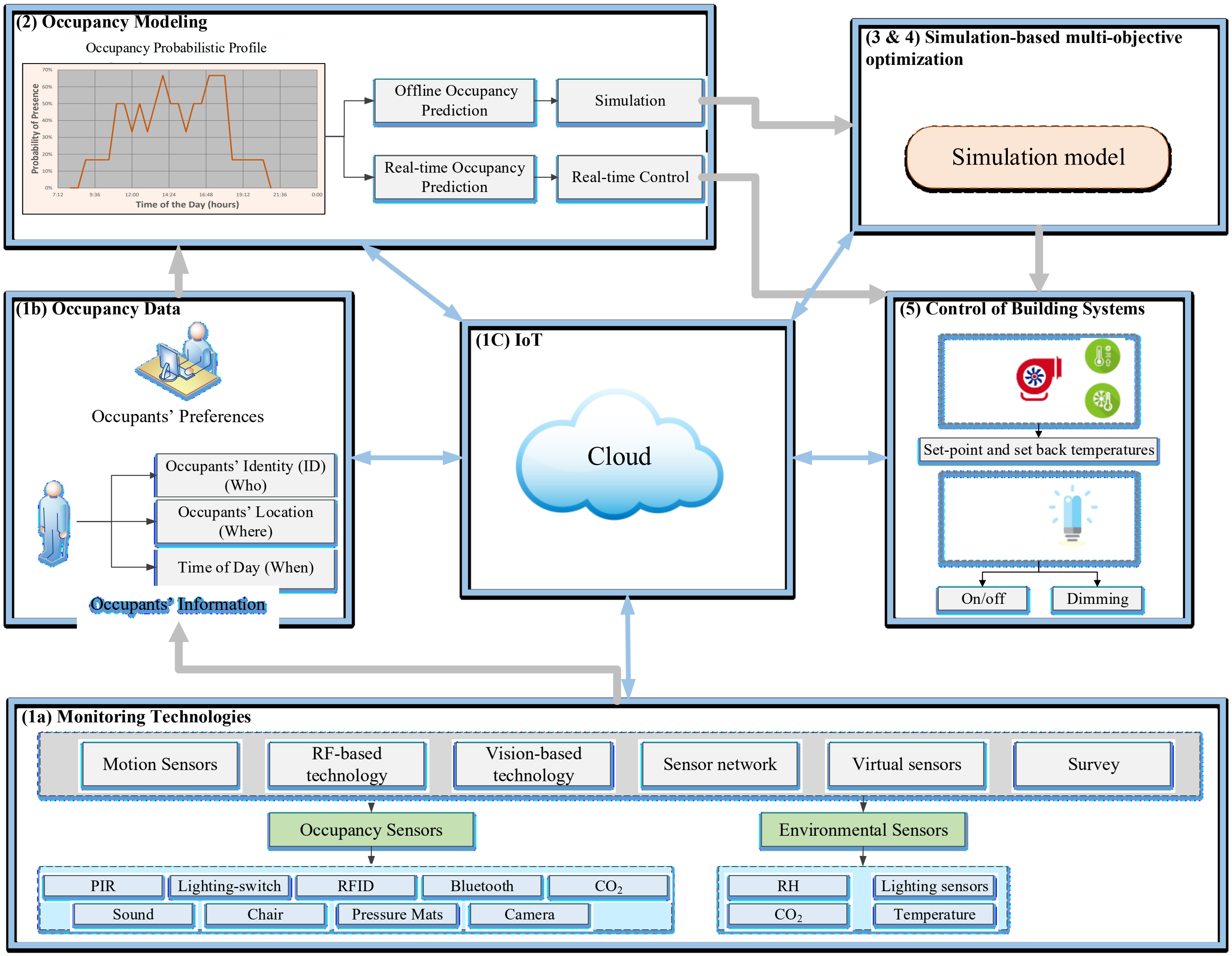


Figure 1-1 Different dimensions of building energy efficiency research related to the occupants' behavior

From the optimization point of view, a limited number of research studies have been conducted that used optimization algorithms to improve the performance of energy conservation methods. Customized occupancy profiles, however, were not included in most of those studies. The studies which used occupancy profiles did not investigate the impact of changes in these profiles on the operation of the building's systems. The effect of occupants' interactions in open-plan offices is also missing in the literature. Therefore, there is a need to integrate the optimization algorithm with building simulation on the basis of dynamic occupancy patterns for open-plan offices.

1.3 Research Objectives

Given the problem definition in Section 1.2, the main objectives of this research are defined as follows:

1. Developing a method for extracting detailed occupancy information with varying time-steps from collected RTLS occupancy data. This method captures different resolution levels required for the application of intelligent, occupancy-centered local control strategies of different building systems.
2. Developing a new time-dependent inhomogeneous Markov chain occupancy prediction model based on the derived occupancy information, which distinguishes the temporal behavior of different occupants within an open-plan office.
3. Improving the performance of the developed occupancy prediction model by determining the near-optimum length of the data collection period, selecting the near-optimum training dataset, and finding the most satisfying temporal resolution level for analyzing the occupancy data.
4. Developing local control algorithms for building energy-consuming systems.
5. Integrating the energy simulation model of an open-plan office with an optimization algorithm to optimally control the building energy-consuming systems and to analyze the trade-off between building energy consumption and occupants' comfort.

These objectives aim at assisting decision-makers in evaluating optimized occupancy-centered building operations and investigating the effect of their application on building energy performance.

1.4 Thesis Organization

The structure of the thesis is as follows:

Chapter 2 Literature Review: In this chapter, a critical review is provided to (1) review different monitoring techniques used to collect occupancy data; (2) identify the detailed aspects of occupancy modelling and the research gaps in each aspect; (3) encapsulate the effect of different occupancy information on the application of integrated building systems control strategies; (4) recognize the research gaps in each; and (5) propose a roadmap regarding the advances in different dimensions of BEMS, including reliable monitoring and data collection techniques, occupancy modeling, and building operational systems control strategies in cognitive buildings.

Chapter 3 Research Framework: The overview of this research and the overall proposed framework are discussed briefly in this chapter. It includes the explanation of the two main modules that are used in the research.

Chapter 4 Probabilistic Occupancy Prediction Model: In this chapter, the occupancy modeling (i.e., occupants' profiles) has been further enhanced using an inhomogeneous Markov chain prediction model, which differentiates the temporal behavior of occupants within an open-plan office based on occupancy space utilization patterns data. Moreover, the required inputs to the model are identified. To provide the validation of the applicability of the proposed model, the building's real occupancy is compared with the results obtained from the prediction model.

Chapter 5 Sensitivity Analyses of the Occupancy Prediction Model: The performance of the proposed prediction model is evaluated in this chapter to find the most effective data collection period and resolution level, which helps the prediction model to produce reliable occupancy information.

Chapter 6 Simulation-based Multi-objective Optimization: This chapter discusses the research methodology employed to produce local control strategies. To do so, a simulation-based multi-objective optimization model is proposed. After creating a simulation model, an optimization algorithm is designed to satisfy the two objective functions of minimizing the office building's energy consumption and occupants' discomfort. The implementation and applicability of the proposed framework are demonstrated through a case study.

Chapter 7 Summary, Conclusions, and Future Work: In this chapter, a summary of this research study is presented and its contributions are highlighted. Moreover, the limitations of the current work are investigated and finally, the recommendations for future research are suggested.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

The fact that buildings contribute a large portion of the global energy consumption arises a need for detailed investigation towards more effective energy performance of buildings. Thus, monitoring, estimating, and reducing buildings' energy consumption have always been important concerns for researchers and practitioners in the field of energy management. Therefore, the intelligent use of energy within buildings is a recent trend of research studies and is the target of Building Energy and Comfort Management (BECM) systems (Nguyen and Aiello, 2013). Figure 2-1 shows the breakdown of energy-consuming systems in commercial/institutional buildings in Canada. With accounting for 71% of total energy use for space heating-cooling and lighting, these systems are the prime targets for energy consumption optimization in order to have a more realistic estimation of buildings' operational energy consumption (NRCAN, 2016).

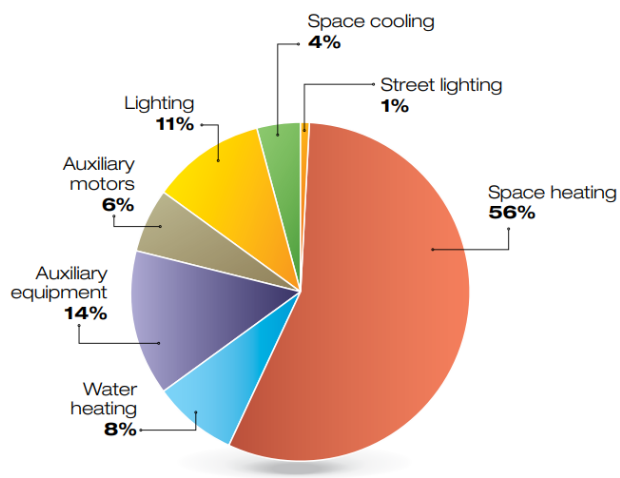


Figure 2-1 Commercial and institutional building energy use, 2014 (NRCAN, 2016)

Many research studies investigated the most important factors affecting the buildings' energy consumption. According to Yu et al. (2011), the factors influencing the total building energy consumption can be divided into seven categories:

(1) Climate (e.g., outdoor air temperature, solar radiation, wind velocity, etc.),
(2) Building-related characteristics (e.g., type, area, orientation, etc.)
(3) User-related characteristics, except for social and economic factors (e.g., user presence, etc.),

(4) Building services systems and operation (e.g., space cooling/ heating, hot water supplying, etc.),
(5) Building occupants' behavior and activities,
(6) Social and economic factors (e.g., degree of education, energy cost, etc.), and
(7) Indoor environmental quality required.

Among these factors, the last two categories represent occupants' influences that affect building energy consumption indirectly and their effect is considered within the study of the fifth category. Occupants' interactions include settings of indoor thermal, acoustic, and visual comfort criteria, opening/closing windows, turning on/off or dimming lights, turning on/off office equipment, and turning on/off the HVAC system. Therefore, building occupants' preferences and activities could have high positive or negative impacts on energy conservation (Yu et al., 2011). These effects are driving factors causing large discrepancies in building energy usage even between similar buildings with the same characteristics and located at similar locations.

There are basically three main sectors investigated by researchers from the energy efficiency point of view including commercial, residential, and other sectors. Offices constitute the largest portion of the commercial sector and are the focus of this research, which is limited to offices within university buildings. The main energy-consuming systems in offices are the HVAC and lighting systems, which are responsible for 33% and 25% of the total energy consumption, respectively (Nguyen and Aiello, 2013). Further, since the application of occupancy detection systems and occupancy-based control strategies differ based on the nature of the building (e.g., residential vs. commercial buildings), the focus of this dissertation is only on office buildings especially open-plan offices.

As a result of improper usage of buildings' devices and systems by their occupants, many buildings actually spend more on energy than is necessary. Computers are kept on while disabling the power saving mode all day long even when occupants left their offices, lights are often left on to illuminate unoccupied rooms, or set to produce more light than necessary. HVAC systems are drift from their ideal settings for efficient performance to satisfy occupants' preferences and are set based on the peak occupancy regardless of actual room usage. Therefore, occupants' presence and preferences have an important effect on the buildings' energy consumption, which should be considered as accurate as possible when dealing with the buildings' energy usage models.

This chapter aims to provide a critical review and a research roadmap of office building energy management based on occupancy monitoring. The main objectives of this chapter are: (1) reviewing different monitoring techniques used to collect occupancy data; (2) identifying the detailed aspects of occupancy modelling; (3) encapsulating the effect of different occupancy information on the application of integrated building systems control strategies; (4) recognizing the research gaps in each dimension and providing linkage between these dimensions and adding structure to them; and (5) proposing a roadmap regarding the advances in different dimensions of BEMS, including reliable monitoring and data collection techniques, occupancy modeling, and building operational systems control strategies in cognitive buildings. The objective of the proposed roadmap is to develop an insight towards the future of Information and Communication Technologies (ICT) in the building sector. The proposed roadmap focuses on emerging ICT and opportunities for the building industry.

2.2 Occupancy Monitoring

To analyze and predict occupants' profiles, occupants should be monitored over a long period of time and occupancy data should be collected. In this context, the occupancy data are mainly categorized into two groups including the data related to the occupants' presence and location and the data regarding the occupants' preferences and their interactions with various energy-related systems within buildings. To collect the first type of data, presence detection systems, such as motion sensors, are used to determine if an occupant is present in a space. However, the exact position of the occupant is still unknown. To find the location of an occupant (i.e., the x and y coordinates), a Real-time Location System (RTLS) can be utilized. These are wireless tracking systems that automatically identify and track the location of objects or people in a defined space in near-real-time (Ward, 2007). Examples of RTLSs are vision-based systems (e.g., cameras), and radiofrequency (RF)-based systems (Hightower and Borriello, 2001).

On the other hand, surveys are usually used to identify the occupants' preferences and the most influential factors that affect the way occupants interact with building systems, such as windows, HVAC, lighting, blinds, and electrical equipment. Using surveys helps to collect information about the occupants' preferences related to the settings of these systems, the occupants' energy-related decisions and their interactions with building systems. Internal personal visual surveys, such as building walkthroughs, are also used to gather data about the building occupants.

These technologies and methods are discussed in more detail in Sections 2.2.1-2.2.6. Furthermore, Table 2-1 summarizes the main research papers related to occupancy monitoring methods, along with different types of sensors used by these methods.

2.2.1 Motion Sensors

Motion sensors are widely used to detect the movement of occupants and to obtain binary occupancy data (i.e., whether an occupant is present in a specific space or not). Ultrasonic detectors, vibration, and infrared (e.g., passive infrared (PIR)) systems, pressure sensors attached to chairs, and magnetic-based approaches (e.g., inertial measurement units (IMUs)) are some examples of motion sensors.

Labeodan et al. (2015, 2016) compared the performance of mechanical-switch sensors (called "chair sensors") with those of strain and vibration sensors in detecting occupancy in open-plan offices. They collected the occupancy data for three days and found that mechanical-switch sensors have the best performance with 99% accuracy followed by the strain and vibration sensors with 95% and 87% accuracy, respectively.

Despite the popularity of motion sensors, they suffer from some fundamental drawbacks especially when it comes to detecting occupancy in a shared space. The first shortcoming is regarding the tracking technique used by motion sensors to collect the occupancy data. For instance, PIR sensors work based on the change in the temperature pattern across the field of view of the sensor, which indicates the presence of an object. Thus, in order to get reliable occupancy information, occupants should be in the field of view of the sensors (Guo et al., 2010). Other types of motion sensors that do not require a field of view to detect occupancy, such as ultrasonic detectors, are prone to be triggered by false movements, such as an occupant moving in an adjacent room. These systems emit ultrasound pulses to detect occupants' movement and any interruption in the transmitted pulse indicates the presence of an occupant. Thus, any false movement can cause errors in detecting occupants. These errors are called false-positive errors and result in conditioning space while it is unoccupied (Harris and Cahill, 2005). Furthermore, professional tuning and commissioning are required to reach a good performance of motion sensors; otherwise, a big percentage of these sensors (more than half) may not work according to the manufacturers' claims regarding their coverage capacity (NLPIP, 1998). Professional tuning and commissioning include changing the positions of sensors, adjusting their angles, and sensitivity tuning (Guo et al., 2010).

Based on the above discussion, motion sensors detect occupancy in single-occupied offices with high accuracy if installed in the right position. However, when they are used in open-plan offices, they are unable to provide detailed occupancy data, such as the number of occupants, their identities, and their activities (i.e., working at their stations, working in other parts of the space, and leaving the space). In addition, the need for a large number of sensors to cover large spaces makes their implementation very costly compared to RF-based systems (Hightower and Borriello, 2001; Pradhan et al., 2009).

Table 2-1 Categorization of research works based on different occupancy monitoring systems (77 papers)

Monitoring Method	Types of Sensors	Example References
Motion sensor	NS*	Wang et al. (2005); Page et al. (2008); Yu (2010)
	PIR	Dodier et al. (2006); Duarte et al. (2013)
	Lighting-switch sensors	Jazizadeh and Becerik-Gerber (2012); Chang and Hong (2013)
	Pressure sensors	Labeodan et al. (2015, 2016)
	Ultrasonic sensors	Harle and Hopper (2008)
Vision-based technology	Camera	Benezeth et al. (2011); Shih (2014); Wang and Ding (2015); Chen et al. (2015)
	Image-processing occupancy sensor	Brackney et al. (2012)
RF-based technology	RFID	Zhen et al. (2008); Li et al. (2012)
	Bluetooth	Harris and Cahill (2005); Conte et al. (2014)
	Wi-Fi	Balaji et al. (2013); Chen and Ahn (2014); Jain and Madamopoulos (2016); Wang and Shao (2017a); Wang and Shao (2017b); Wang et al. (2017); Çiftler et al. (2017); Wang et al. (2018); Yang et al. (2018)
Multi-sensor Networks	Cameras, PIR, and CO_2 sensors	Meyn et al. (2009)
	Wired CO_2 and indoor air quality sensing network, wireless ambient sensing network	Lam et al. (2009); Dong et al. (2010)
	Smart camera networks	Erickson and Cerpa (2010); Erickson et al. (2010, 2011); Cho et al. (2010)
	Contact closure, PIR, and CO_2 sensors	Newsham and Birt (2010); Dedesko et al. (2015)
	PIR, CO_2, sound, light, and power use sensors	Hailemariam et al. (2011)
	PIR, CO_2, RH, light, and temperature sensors	Attar et al. (2011)
	RFID, RH, light, and temperature sensors	Augello et al. (2011)
	PIR, CO_2, sound, light, and temperature sensors	Yang et al. (2012)
	PIR, CO_2, RH, temperature, air velocity and globe thermometer	Han et al. (2012)
	PIR, pressure mats, personal computers, CO_2, VOC, temperature, RH, acoustics, light dependent resistor (LDR)	Ekwevugbe et al. (2012)
	PIR, pressure, and acoustic sensors	Nguyen and Aiello (2012)
	PIR, ultrasound sensors and power plug meters	Milenkovic and Amft (2013)

*NS: Not specified in paper

Table 2-1 Categorization of research works based on different occupancy monitoring systems (Cont.)

Monitoring Method	Types of Sensors	Example References
Multi-sensor Networks	RH, temperature, CO_2, VOC, motion, and light sensors	Fabi et al. (2014)
	Light, temperature, humidity, audio level, PIR sensors, meeting schedules, and computer activity	Khan et al. (2014)
	Ultrasound, pressure sensors, Wi-Fi, and power meters	Jin et al. (2014)
	PIR, CO_2, temperature, RH, air-velocity sensors, global thermometer, and reed switches	Ai et al. (2014)
	Camera, light, temperature, RH, PIR, door contact, CO_2, and power meters	Arora et al. (2015)
	Smart Door (LDR and ultrasonic Sensors)	Nasir et al. (2015)
	Wi-Fi and light sensors	Mohammadmoradi et al. (2017)
	PIR, CO_2, VOC, temperature, RH, acoustics, and light sensors, and camera)	Ekwevugbe et al. (2017)
	Keyboard and mouse activity, webcam, microphone, PIR, temperature, RH, light, proximity sensors, and pressure mat	Newsham et al. (2017)
	Temperature, humidity, light, and CO_2 sensors	Nesa and Banerjee (2017)
	CO_2, magnetic reed switches, and PIR sensors	Javed et al. (2016)
	IMU, Wi-Fi, humidity, and illuminance sensors	Zhao et al. (2017)
	Wi-Fi and BLE	Mashuk et al. (2018)
Virtual sensors	PIR, pressure, and keyboard and mouse sensors, GPS location and Wi-Fi connection from Wi-Fi hotspots	Zhao et al. (2015); Jin and Spanos (2017)
Survey		Brager et al. (2004); Karjalainen (2007); Tabak (2008); Tabak and de Vries (2010); Liao and Barooah (2010); Wei et al. (2010); Goldstein et al. (2010a, 2010b); Goldstein et al. (2011); Haldi and Robinson (2008, 2011); Kavulya and Becerik-Gerber (2012); Balaji et al. (2013); Purdon et al. (2013); Humphreys et al. (2013); Fabi et al. (2014); Sun et al. (2014); West et al. (2014); Day and Gunderson (2015); Wang and Ding (2015); Hong et al. (2015a, 2015b)

*NS: Not specified in the paper

2.2.2 Vision-Based Localization Technologies

To alleviate the shortcomings of motion sensors, vision, and RF-based localization technologies were introduced to distinguish between different occupants and track them according to their identities. This information provides better insight into the usage pattern of shared spaces. Benezeth et al. (2011) presented a vision-based algorithm to capture detailed occupancy information by combining background subtraction, tracking, and recognition. They utilized static cameras to acquire information regarding occupant presence, location, number, and types of activities. The proposed method provided information on the presence or absence of occupants with 97% accuracy. Although vision-based systems have a high detection rate, the privacy concern (an area of increasing interest) and the heavy image processing steps (required to extract occupancy data) restrict their wide implementation.

2.2.3 RF-Based Localization Technologies

RF-based localization technologies use radio frequency to position an object and include ultra-wideband (UWB), radio-frequency identification (RFID), Wireless Local Area Network (WLAN) or Wi-Fi, Global System for Mobile communication (GSM), Bluetooth, and ZigBee. Recently, these systems became more popular due to their deployment flexibility, communication range and ability to work without a line of sight (Soltani et al., 2015). A system of multiple active RFID readers was implemented by Zhen et al. (2008) to determine the occupant location in an indoor environment. They also used multiple readers to handle the multipath effect of RFID. Therefore, instead of having received signal strength indicator (RSSI) by one reader, the system recorded an RSSI vector that represents the RSSI by multiple readers. The proposed localization system showed lower accuracy for stationary occupants compared to mobile ones due to the usage of multiple readers and the RSSI vector.

Li et al. (2012) proposed an RFID based occupancy detection system to control the operation of HVAC systems. The system detects and tracks stationary and mobile occupants in multiple single- and multi-occupied spaces. Thus, it detects the location of each occupant and the number of occupants in each thermal zone. The system can detect occupants at the zone level with an accuracy of 88% for stationary occupants and 62% for mobile occupants. By testing the operation methods and determining the major energy consumers in HVAC systems, they proposed eight energy-saving strategies. However, they did not test the efficiency of the proposed strategies in the field study. Their research showed that using RFID in small spaces does not result in promising occupancy detection due to the signal interference of occupants' tags. In addition, reference tags' locations are another important factor affecting occupancy detection accuracy. Unlike Zhen et al. (2008), they found that the proposed occupancy detection system using RFID worked better for stationary occupants rather than mobile ones.

The usage of wireless sensing technologies, such as Wi-Fi, has increased in recent years. Almost all modern buildings are equipped with Wi-Fi access points (APs) and wireless devices, making occupancy detection more efficient, affordable, and convenient (Wang et al., 2017). Wi-Fi enabled devices (e.g., laptops, smartphones, and tablets) allow occupants to connect to Wi-Fi networks. This connection then can be used as an occupancy indicator of space. Many studies tracked occupancy using Wi-Fi networks and used the results of occupancy detection for demand-driven control of building systems (Mashuk et al., 2018; Wang et al., 2018). Wang and Shao (2017b) used a Wi-Fi-based indoor positioning system and created occupancy profiles based on the measured Wi-Fi devices' number and locations in a university library building. By assessing the implication of the occupancy patterns for lighting system energy efficiency, they reached 26.1% decrease in the total energy consumption. Wang et al. (2017) determined occupancy with 80% accuracy. Wang et al. (2018) proposed a ventilation strategy based on the detected occupancy profiles using a Wi-Fi probe that enabled an occupancy sensing system. Conducting a two-day experiment in a graduate students' office resulted in ventilation energy consumption savings of 44.26% (weekday) and 55.5% (weekend day) when compared to the fixed-rate ventilation strategy.

In 1998, Bluetooth Special Interest Group (SIG) formalized one of the first standardized wireless technologies using Bluetooth. However, the high energy consumption of transceiver ships, long connection latency, large memory allocation due to a complex protocol stack, and overhead due to large data packets restricted the application of Bluetooth. To resolve these drawbacks, Bluetooth Low Energy (BLE) (or Bluetooth Smart) was launched in 2010 as one of the protocols in the Bluetooth Core Specification version 4.0. The main advantage of BLE is the reduction in energy consumption (up to three years on a single coin battery) and cost (i.e., 60-80% cheaper) while providing a higher communication range than traditional Bluetooth (Mackensen et al., 2012). The broadcast range of BLE is up to 100 meters, which is much further than the classic Bluetooth (10 meters), making BLE perfect for indoor location tracking and awareness (IndustryARC, 2016; Bluetooth, 2016). Considering the wide range of BLE applicability, it is predicted that the BLE-enabled device shipments will increase from 1.8 billion units in 2014 to 8.4 billion units by 2020 (a compound annual growth rate (CAGR) of 29%). Smartphones, tablets, BLE-enabled sensors or any device implementing the BLE standard can be used as a BLE hub. The ability to track multiple moving objects in real-time makes BLE systems optimal RTLSs for different applications, such as building energy efficiency, sport, and healthcare applications, optimizing store layout, security, and emergency situations (Quuppa, 2017).

Harris and Cahill (2005) introduced a context-aware power management (CAPM) system to minimize the electricity consumption of desktop computers. After conducting experimental trials using Bluetooth-enabled mobile phones to detect the occupants' location, they found that location alone is insufficient for effective power management. Thus, they used a Bayesian network to add acoustic sensor data and time of day to the location data. These data provide much information pertinent to occupant behavior, such as the sound of opening/closing doors. Although they mentioned that the proposed model is reliable in personal and shared spaces, no real-world experiments were conducted to show the effectiveness of the method. The proposed method may not efficiently respond to other control strategies, such as lighting controls since it was hard for the BLE system to entirely cover a space. Not considering occupants' identities as well as not associating occupants with rooms may also result in false-positive errors. As mentioned in Section 2.2.1, activating the lighting system for the condition when an occupant is in an adjacent room is an example of the false-positive errors.

After releasing a technology called *iBeacon* by Apple in 2013, this protocol was modified by Conte et al. (2014) based on BLE to be used as an occupancy detection system. They proposed a solution called *BLUE-SENTINEL* to determine the number of occupants, their location and identities using occupants' mobile devices as the data collection system. Implementing the proposed approach in three laboratory rooms showed 83% accuracy.

2.2.4 Multi-sensor Networks

The information coming from only one source of data may be unreliable for occupancy detection. For instance, most of the current sensing technologies (e.g., motion sensors) are unable to determine detailed occupancy information, such as the number of occupants taking up space. To solve this problem, sensor networks are used by many researchers. These networks combine different monitoring technologies to take full advantage of the strong points of their integration and to overcome their limitations when used alone. In multi-sensor networks, occupancy and environmental data are collected from different types of sensors. The data fusion techniques are then applied to fuse the redundant data, select the important features indicating the occupancy in space, and derive the parameters of importance (Krishnamachari et al., 2002). For instance, Ekwevugbe et al. (2012) used a sensor fusion model based on the Adaptive Neuro-Fuzzy Inference System (ANFIS) algorithm to estimate a reliable occupancy profile using data collected from a multi-sensors network.

Meyn et al. (2009) introduced the sensor-utility-network (SUN) system in their experiments using multiple sensors from three classes: (1) 10 digital video cameras; (2) 12 PIR sensors; and (3) 15 CO_2 sensors. This strategy provides more informative data for occupancy estimation at zone and building levels. Compared to the ground truth occupancy obtained manually by analyzing individual video frames, the SUN system estimates the number of occupants at the building level with the estimation error of 11%. The necessity of more accurate occupancy estimations is felt when optimizing the operation of HVAC and lighting systems. Ground truth obtained manually by analyzing individual video frames and correcting them.

A wireless sensor network comprising of contact closure, PIR, and CO_2 sensors were used in the research of Newsham and Birt (2010). They conducted a test in an office building including laboratories and individual workspaces to count the number of occupants. The power demand of the building can then be forecasted using the gathered data. They found that using other types of sensors other than motion sensors increased the accuracy of the prediction model.

Diaz et al. (2011) used wireless sensor networks to monitor the energy consumption of all devices in an intelligent building. Temperature, humidity, luminosity, electrical consumption, and presence sensors were used in the ECoSence project. Their goal was to use the obtained data to improve energy consumption and render the buildings environmentally sustainable.

To detect indoor occupants' activities in a single-occupied office, Nguyen and Aiello (2012) used a simple sensor wireless network (i.e., infrared, pressure, and acoustic). Five activities (i.e., working at a desk with or without a PC, participating in a meeting, the presence, and absence) were recognized by using their prototype while the users' privacy was unaffected as information was recorded in a binary manner (i.e., TRUE, FALSE). These activities can be used as inputs for applying different control strategies.

Erickson et al. (2009) used SCOPES, a distributed Smart Camera Object Position Estimation System proposed by Kamthe et al. (2009), to gather near real-time occupant movement with 80% accuracy in a large multi-function building. Based on the collected data, occupant mobility patterns were predicted by applying Gaussian and agent-based models. They achieved 14% energy savings on the HVAC system by applying an optimal control strategy based on occupants' activity estimates. They found that the Gaussian model performs better for real-time prediction, while the agent-based model results in more energy-efficient building designs. They also examined the performance of the smart camera network using Markov chain models. A 20% annual energy savings was achieved using Markov models for the occupant activity estimates (Erickson and

Cerpa, 2010). In another study, they installed their camera network only on the ceilings of corridors to predict occupancy and reached 42% annual energy savings while maintaining comfort standards (Erickson et al., 2011). Considering the transition points placed at entrances/exits provides no indication of the occupancy patterns and interactions within other building spaces, such as offices, labs, meeting rooms, etc. According to their 48-hour observations, they assumed zero occupancies during nighttime. This assumption cannot be applied in buildings with a different function, such as hospitals, which restricts the usage of the proposed model to specific types of buildings. In addition, when a camera sensor network is used to detect occupancy, many pre- and post-processing algorithms are required to extract the desired data from the collected data. Also, due to the excessive labor required to gather data over long periods of time using this type of sensor network, occupancy data is only collected for 48 hours. In order to increase the efficiency of the model, they considered several assumptions related to the maximum number of people who can move through doorways as well as concurrent movement through several doorways.

Khan et al. (2014) presented a wireless sensor network (WSN) including light, temperature, humidity, audio level and PIR sensors that only collect non-sensitive data by explicitly avoiding privacy-violating means, such as cameras or microphones along with external data sources (i.e., computer activity and meeting schedules). They introduced a hierarchical analysis framework to predict occupancy at three different levels of granularity: (1) binary detection, (2) categorical occupancy estimation, and (3) counting the exact number of occupants. Using statistical classifiers adds confidence levels to different granularity levels. This helps decision-makers to make more reasonable decisions when less detailed, but more reliable information is available. They deployed the proposed framework in a real-world test by monitoring a high-traffic area (i.e., an open-plan office with 20 occupants) and a low-traffic area (i.e., small meeting room) for 10 and 14 days, respectively. The contextual data was however used only for the meeting room. The results demonstrated that the meeting room was not utilized according to schedules for nearly a third of the time. The proposed framework cannot detect the number of occupants with high confidence in high-traffic areas. In addition, the proposed methodology is based on a certain number of occupants (i.e., 14 occupants). Thus, an open-ended occupancy classification problem, where the maximum number of occupants is not strictly pre-defined, is necessary and more practical for real-world applications.

2.2.5 Virtual Occupancy Sensors

Some scholars argue for the cost-effectiveness of special-purpose occupancy sensors, such as motion sensors or vision-based systems. These sensors require setup and commissioning, calibration, and frequent maintenance during their useful life. This makes their application costly especially in the case of a sensor network in large areas. Therefore, virtual occupancy sensors are introduced to provide a non-intrusive and cost-effective way to detect occupants' presence using existing energy-related systems within buildings. For instance, desktop activity and energy meters can be used to provide an indication of occupants' presence in an office. Smart power meters were used by Jin and Spanos (2017) to detect occupancy. Implementation of the proposed method in residential and commercial buildings showed 78-93% and 90% accuracy for residences and offices, respectively. However, using virtual occupancy sensors provides only binary occupancy data (i.e., presence and absence) without indicating other important information, such as occupants' identities and activities. For example, occupants may be present in their office but not using any electrical devices. In this case, no occupancy is reported by virtual occupancy sensors.

To overcome this limitation, the virtual occupancy sensors data are combined with the data of physical occupancy sensors in some applications to derive more accurate occupancy information. Two types of virtual occupancy sensors were introduced by Zhao et al. (2015) at room- and working zone-levels. PIR, pressure, keyboard, and mouse sensors were used for the room-level virtual sensors. Zone-level occupancy detection was performed using a real-time global positioning system (GPS) location and Wi-Fi connection to Wi-Fi hotspots. They integrated all these occupancy measurements using a Bayesian Belief Network (BBN). The performance of the proposed virtual sensor was evaluated by collecting data for one to two weeks from two private offices. The results showed better performance of the combined system than individual use of sensors. They indicated that it would be more convenient to use smart devices, such as smartphones, to get Wi-Fi information. However, battery usage and privacy concerns are the main issues when using personal devices. In addition, the application of the proposed system had not been investigated for more random occupancy patterns and multiple occupants in open-plan offices.

2.2.6 Survey

As mentioned in Section 2, some researchers use surveys as a way of collecting occupant information, either combined with other tracking technologies or alone. For instance, occupants' behavior in five single-occupied offices (regarding their usage pattern of office devices) was investigated by Kavulya and Becerik-Gerber (2012) using in-person observation. In addition, non-intrusive appliance load monitoring was used to track device energy consumption in these offices. They found that a 38% energy savings could be obtained by simply turning off the office devices when they are not in use. To find the potential savings, they used average values of the results derived from all offices. Thus, they did not consider the difference between occupant preferences and their working states. A stochastic occupancy model was developed by Wang and Ding (2015) based on the correlation between the occupants' activities and equipment energy consumption. They used cameras to monitor the occupants' behavior and count the number of occupants in each time step. They also conducted a survey to gather data about occupants' working habits to determine different absence states.

A model-free HVAC control algorithm was proposed by Purdon et al. (2013) to avoid installing sensors or building complex occupant comfort models. The control algorithm considers occupant preferences regarding the HVAC system settings through an application, which collects their votes. Due to changes in the office occupancy from day to day and the complexity of implementing different control strategies in a real case, they evaluated the performance of the proposed control algorithm through a simulator. The simulator used empirical data from 20 occupants located in 12 offices and one conference room and simulated different HVAC control strategies. They created the comfort models of eight participants by specifying their comfort limits based on filtering for outliers in the information gathered from the surveys. A high degree of correlation between the preferences of individual participants suggested that it is possible to reach an internal temperature that makes most participants comfortable.

2.3 Occupancy Modeling

Occupancy models are developed using the data collected during the occupancy monitoring period. These models could then predict the probability of occupancy and various occupant activities under different conditions. Tracking data provide insights to different occupancy information, such as the number of occupants, their location, and their identities for each space (and each time step) of a building. Moreover, occupants interact with office buildings in different ways. They work in their offices that may be private or open-plan offices. They communicate with their colleagues and other occupants in other spaces of the building. Also, they occasionally gather in meeting rooms. After collecting the occupancy data, analysis is required to determine occupant activity and other occasional variations in occupant schedules.

Currently, most of occupancy schedules used in building energy simulation are considered binary and deterministic, with the Boolean values of '0' and '1' representing unoccupied and occupied states of the space, respectively. Although some diversity can be considered by using different deterministic schedules for workdays and weekends, all workdays are considered to have the same profile throughout the year (Davis and Nutter, 2010). This results in the same level of energy consumption in all spaces within the building. Furthermore, using simplified deterministic schedules in the building simulation results in a discrepancy between the building's actual energy consumption and the results of the energy simulation. That is due to the inability of the deterministic schedules to consider the variations of the energy consumption in the cases of special events. Also, the peak load of spaces may be overestimated, as these schedules consider the maximum occupancy in all spaces at the same time. However, this situation rarely happens in office buildings. Thus, more precise and detailed occupancy models should be integrated with simulation tools to more realistically estimate energy consumption of buildings.

There are basically three types of methods (statistical, stochastic, and ML) that represent the probabilistic occupancy models in this study. These methods and related research works are explained in Sections 2.3.1-2.3.3. Furthermore, Table 2-2 summarizes the comparison of the research works that have been reviewed, with an emphasis on the data collected by means of different occupancy monitoring methods to create probabilistic occupancy models. In the context of this research, location, as shown in Table 2-2, refers to the x and y coordinates of the occupant(s). Thus, not marking the location for a paper implies that the monitoring method only detects the presence of the occupant(s) at room/space level and not the exact location of the occupant(s). Number, identity, duration, and activity data provide answers to the following questions: (1) How many people are present in a space? (2) Who are they? (3) For how long is/are the occupant(s) present in the space? (4) What is/are the occupant(s) doing in the space (i.e., working at their stations, working in other parts of the space, and leaving the space)? However, since all the references in Table 2-2 considered the duration parameter, this parameter is not shown in this table to avoid repetition.

In addition, Table 2-3 categorized these studies based on the type of the analysis method they have used to develop the probabilistic occupancy models. This table shows the list of occupancy modeling methods that are widely used by scholars for each of the above-mentioned categories. The research studies are also compared based on the type of study and spaces used as demonstrated in Table 2-4. More than half of the reviewed papers provided binary occupancy information and the rest emphasized on counting the number of occupants. In addition, almost half of them determined occupants' activities. Among the 80 papers reviewed, only seven detected occupancy at the identity level and only 20 provided detailed information regarding occupant location. It can be seen from Table 2-4 that simulation was used only by 20 studies and only two papers applied optimization to detect occupancy behavior based on behavioral rules.

Table 2-2 Comparison of research papers focusing on probabilistic occupancy modeling using monitoring technologies (80 papers)

Reference	Occupancy Monitoring	Occupancy Model Resolution			
		Location	**Number**	**Identity**	**Activity**
Yamaguchi et al. (2003)	-	-	-	-	✓
Brager et al. (2004)	Survey	-	-	-	✓
Karjalainen (Karjalainen, 2007)	Survey	-	-	-	✓
Wang et al. (2005)	Motion sensor	-	-	-	-
Harris and Cahill (2005)	Bluetooth-enabled mobile phones, acoustic sensors	✓	-	-	✓
Dodier et al. (2006)	PIR sensors	-	-	-	-
Page et al. (2008)	Motion sensor	-	-	-	-
Harle and Hopper (2008)	Ultrasonic sensors (CAMP)	✓	-	-	-
Zhen et al. (2008)	RFID	✓	✓	-	-
Tabak (2008)	A web-based survey	✓	-	-	✓
Haldi and Robinson (2008)	Survey and temperature sensors	-	-	-	✓
Meyn et al. (2009)	Camera, PIR, and CO_2 sensors	-	✓	-	-
Lam et al. (2009)	Wired CO_2 and indoor air quality sensing network (CO_2, CO, TVOC, temperature), wireless ambient sensing network (PIR, RH, sound sensors)	-	✓	-	✓
Tabak and de Vries (2010)	Survey	-	-	-	✓
Liao and Barooah (2010)	Motion sensor	-	-	-	-
Daum and Morel (2010)	Motion sensor	-	-	-	-
Dong et al. (2010)	Wired CO_2 and indoor air quality sensing network (CO_2, CO, TVOC, temperature), wireless ambient sensing network (PIR, RH, sound sensors)	-	✓	-	✓
Cho et al. (2010)	Smart camera networks	✓	✓	-	-
Newsham and Birt (2010)	Contact closure, PIR, and CO_2 sensors	-	✓	-	-
Yu (2010)	Motion sensor	-	-	-	✓
Wei et al. (2010)	Survey	-	✓	-	✓
Goldstein et al. (2010a, 2010b)	Survey	-	✓	-	✓
Erickson and Cerpa (2010) ; Erickson et al. (2010, 2011)	Smart camera networks	✓	✓	-	✓
Goldstein et al. (2011)	Survey	✓	✓	-	✓
Wang et al. (2011)	-	✓	✓	-	-
Benezeth et al. (2011)	Static cameras	✓	✓	✓	✓
Hailemariam et al. (2011)	PIR motion sensor, CO_2, sound, light, and power use sensors	-	-	-	-
Augello et al. (2011)	RFID, RH, light, and temperature sensors	-	✓	✓	✓
Attar et al. (2011)	PIR, CO_2, RH, light, and temperature sensors	-	✓	-	-
Virote and Neves-Silva (2012)	Visual observation	-	-	-	✓

Table 2-2 Comparison of research papers focusing on probabilistic occupancy modeling using monitoring technologies (Cont.)

Reference	Occupancy Monitoring	Occupancy Model Resolution			
		Location	Number	Identity	Activity
Ekwevugbe et al. (2012)	PIR, pressure mats, personal computers, CO_2, VOC, temperature, RH, acoustics, and LDR	-	✓	-	✓
Nguyen and Aiello (2012)	Infrared, pressure, and acoustic sensors	-	-	-	✓
Kavulya and Becerik-Gerber (2012)	Visual observation, non-intrusive appliance load monitoring	-	-	-	✓
Jazizadeh and Becerik-Gerber (2012)	Light intensity sensors	-	-	-	✓
Brackney et al. (2012)	Image-processing occupancy sensor	-	✓	-	-
Yang et al. (2012)	PIR, CO_2, sound, light, and temperature sensors	-	✓	-	-
Han et al. (2012)	PIR, CO_2, and RH, temperature, air velocity and globe thermometer	-	✓	-	-
Chang and Hong (2013)	Lighting-switch sensors	-	-	-	✓
Duarte et al. (2013)	PIR sensors	-	-	-	-
Milenkovic and Amft (2013)	PIR and power plug meters	-	✓	-	✓
Humphreys et al. (2013)	Survey	-	-	-	✓
Fabi et al. (2014)	Survey, RH, temperature, CO_2, VOC, motion, and light sensors	-	-	-	✓
Sun et al. (2014)	Survey, occupants' access cards	-	✓	✓	-
Conte et al. (2014)	BLUE-SENTINEL beacons	✓	✓	✓	-
Khan et al. (2014)	light, temperature, humidity, audio level, PIR sensors, meeting schedules, and computer activity	-	✓	✓	-
Jin et al. (2014)	Ultrasound, pressure sensors, Wi-Fi, and power meters	-	-	-	-
Chen and Ahn (2014)	Wi-Fi	✓	-	-	-
Shih (2014)	Camera	✓	✓	-	-
Ai et al. (2014)	PIR, CO_2, temperature, RH, air-velocity sensors, global thermometer, and reed switches	-	✓	-	-
Feng et al. (2015)	-	✓	✓	-	-
Chen et al. (2015)	Camera	-	✓	-	-
D'Oca and Hong (2015)	Motion sensor	-	✓	-	-
Dedesko et al. (2015)	CO_2 and doorway IR beam-break sensors	-	✓	-	✓
Wang and Ding (2015)	Camera, survey, and power meters	-	✓	-	✓
Nasir et al. (2015)	Smart Door (LDR and ultrasonic Sensors)	-	✓	✓	-
Day and Gunderson (2015)	Survey	-	-	-	✓
Hong et al. (2015a, 2015b)	Survey	✓	✓	✓	✓
Zhao et al. (2015)	PIR, pressure, and keyboard and mouse sensors, GPS location and Wi-Fi connection from Wi-Fi hotspots	✓	-	-	-
Arora et al. (2015)	Camera, light, temperature, RH, PIR, door contact, CO_2, and power consumption	-	✓	-	-

Table 2-2 Comparison of research papers focusing on probabilistic occupancy modeling using monitoring technologies (Cont.)

Reference	Occupancy Monitoring	Occupancy Model Resolution			
		Location	Number	Identity	Activity
Arora et al. (2015)	Camera, light, temperature, RH, PIR, door contact, CO_2, and power consumption	-	✓	-	-
Labeodan et al. (2015, 2016)	Pressure, strain, vibration, and PIR sensors	-	✓	-	-
Jain and Madamopoulos (2016)	Wi-Fi	✓	✓	-	-
Javed et al. (2016)	CO_2, magnetic reed switches, and PIR sensors	-	✓	-	-
Mohammadmoradi et al. (2017)	Wi-Fi and light sensors	-	✓	-	-
Ekwevugbe et al. (2017)	PIR, CO_2, VOC, temperature, RH, acoustics, and light sensors, and camera	-	✓	-	✓
Newsham et al. (2017)	Keyboard and mouse activity, webcam, microphone, PIR, temperature, RH, light, proximity sensors, and pressure mat	-	-	-	✓
Jin and Spanos (2017)	Commercial: Ultrasonic, acceleration (attached to chair), Wi-Fi and survey Residential: Electricity power meters, manual entry	-	-	-	-
Wang et al. (2017)	Wi-Fi and camera	-	✓	-	-
Wang and Shao (2017a, 2017b)	Wi-Fi and light sensors	✓	✓	-	-
Nesa and Banerjee (2017)	Temperature, humidity, light, and CO_2 sensors	-	-	-	✓
Zhao et al. (2017)	IMU, Wi-Fi, humidity, and illuminance sensors	✓	✓	-	-
Çiftler et al. (2017)	Wi-Fi	✓	✓	-	-
Wang et al. (2018)	Wi-Fi	-	✓	-	-
Yang et al. (2018)	Wi-Fi	-	-	-	✓
Mashuk et al. (2018)	Wi-Fi and BLE	✓	✓	-	-

Table 2-3 Categorization of research papers based on the type of the probabilistic occupancy modeling method (89 papers)

Analysis Method	**Type**		**References**
Statistical	Linear Regression		Wang et al. (2005); Mahdavi et al. (2008); Goldstein et al. (2010a, 2010b); Goldstein et al. (2011); Humphreys et al. (2013)
	Bayesian probability		Harris and Cahill (2005); Dodier et al. (2006); Meyn et al. (2009); Langevin et al. (2013); Zhao et al. (2015); Mashuk et al. (2018)
	Logistic Regression		Tabak (2008); Wang et al. (2005); Tabak and de Vries (2010); Liao and Barooah (2010); Daum and Morel (2010); Haldi and Robinson (2008, 2011); Chang and Hong (2013); Gunay et al. (2014); Fabi et al. (2014)
	Support Vector Regression (SVR)		Wang et al. (2017)
	t-test		Brager et al. (2004); Duarte et al. (2013); Day and Gunderson (2015)
	U test		Karjalainen (2007)
	Pearson chi-square test		Day and Gunderson (2015)
	KS test		Sun et al. (2014)
	Time Series		Feng et al. (2015)
Stochastic	Standard Markov Model		Yamaguchi et al. (2003); Page et al. (2008); Wei et al. (2010); Wang et al. (2011); Erickson et al. (2011); Han et al. (2012); Dong and Lam (2011); Dobbs and Hencey (2014a, 2014b); Chen et al. (2015); Jain and Madamopoulos (2016)
	MCMC		Wang and Ding (2015)
	HMM		Lam et al. (2009); Dong et al. (2010); Dong and Lam (2011); Virote and Neves-Silva (2012); Han et al. (2012)
	Layered Hidden Markov Model (LHMM)		Milenkovic and Amft (2013)
	Autoregressive Hidden Markov Model (ARHMM)		Han et al. (2012); Ai et al. (2014); Wang et al. (2017)
	Dynamic Markov Time-Window Inference (DMTWI)		Wang et al. (2017)
	Various Probability Distributions		Hong et al. (2015a, 2015b)
Machine Learning	SVM		Zhen et al. (2008); Lam et al. (2009); Dong et al. (2010); Shih (2014); Nasir et al. (2015); Jin and Spanos (2017)
	ANN		Lam et al. (2009); Dong et al. (2010); Ekwevugbe et al. (2012); Yang et al. (2012); Javed et al. (2016); Wang et al. (2017); Ekwevugbe et al. (2017)
	Decision Tree		Wei et al. (2010); Hailemariam et al. (2011), D'Oca and Hong (2015); Arora et al. (2015); Newsham et al. (2017); Capozzoli et al. (2017)
	Classification methods		Khan et al. (2014); D'Oca and Hong (2015); Nesa and Banerjee (2017); Zhao et al. (2017); Yang et al. (2018)
	Polynomial Regression (Cubic)		Wang and Ding (2015)
	Clustering	K-means	Augello et al. (2011); D'Oca and Hong (2015); Capozzoli et al. (2017); Wang and Shao (2018)
		K-nearest neighbor (KNN)	Peng et al. (2017)
		Affinity propagation	Jain and Madamopoulos (2016)
	Bayesian networks		Augello et al. (2011)
	PresenceSense (PS)		Jin et al. (2014)
	Optimization (GP)		Yu (2010); Newsham et al. (2017)

Table 2-4 Comparison of research papers based on type of study and space (80 papers)

Reference	Type of Study			Type of Space	
	Optimization	Simulation	Field Study	Shared	Private
Yamaguchi et al. (2003)	-	✓	-	-	✓
Brager et al. (2004)	-	-	✓	✓	✓
Karjalainen (Karjalainen, 2007)	-	✓	✓	✓	-
Wang et al. (2005)	-	-	✓	-	✓
Harris and Cahill (2005)	-	-	-	-	-
Dodier et al. (2006)	-	-	✓	-	✓
Page et al. (2008)	-	-	✓	-	✓
Harle and Hopper (2008)	-	-	✓	✓	-
Zhen et al. (2008)	-	-	✓	-	✓
Tabak (2008)	-	✓	✓	✓	✓
Haldi and Robinson (2008)	-	-	✓	NS*	NS*
Meyn et al. (2009)	-	-	✓	NS*	NS*
Lam et al. (2009)	-	-	✓	✓	-
Tabak and de Vries (2010)	-	-	✓	NS*	NS*
Liao and Barooah (2010)	-	✓	-	✓	✓
Daum and Morel (2010)	✓	✓	-	-	✓
Dong et al. (2010)	-	-	✓	✓	-
Cho et al. (2010)	-	✓	-	✓	-
Newsham and Birt (2010)	-	-	✓	✓	✓
Yu (2010)	✓	-	-	-	✓
Wei et al. (2010)	-	✓	✓	✓	✓
Goldstein et al. (2010a, 2010b, 2011)	-	✓	-	✓	✓
Erickson and Cerpa (2010) ; Erickson et al. (2010, 2011)	-	-	✓	✓	-
Wang et al. (2011)	-	✓	-	✓	✓
Benezeth et al. (2011)	-	-	✓	✓	✓
Hailemariam et al. (2011)	-	-	✓	✓	-
Augello et al. (2011)	-	-	✓	✓	-
Attar et al. (2011)	-	✓	✓	-	✓(Cubical)
Virote and Neves-Silva (2012)	-	✓	-	✓	✓
Ekwevugbe et al. (2012)	-	-	✓	✓	-
Nguyen and Aiello (2012)	-	-	✓	-	✓
Kavulya and Becerik-Gerber (2012)	-	-	✓	-	✓
Jazizadeh and Becerik-Gerber (2012)	-	-	✓	NS*	NS*
Brackney et al. (2012)	-	✓	✓	✓	-
Yang et al. (2012)	-	-	✓	✓	-
Han et al. (2012)	-	-	✓	✓	-
Chang and Hong (2013)	-	-	✓	✓	-
Duarte et al. (2013)	-	-	✓	✓	✓
Milenkovic and Amft (2013)	-	✓	✓	✓	✓
Humphreys et al. (2013)	-	-	✓	NS*	NS*
Fabi et al. (2014)	-	-	✓	✓	✓
Sun et al. (2014)	-	✓	✓	NS*	NS*
Conte et al. (2014)	-	-	✓	✓	-
Khan et al. (2014)	-	-	✓	✓	-
Jin et al. (2014)	-	-	✓	✓	-
Chen and Ahn (2014)	-	-	✓	✓	-
Shih (2014)	-	✓	✓	✓	-
Ai et al. (2014)	-	-	✓	✓	-
Feng et al. (2015)	-	✓	-	✓	✓
Chen et al. (2015)	-	✓	-	✓	-
D'Oca and Hong (2015)	-	-	✓	-	✓

Table 2-4 Comparison of research works based on type of study and space (Cont.)

Reference	Type of Study			Type of Space	
	Optimization	Simulation	Field Study	Shared	Private
Dedesko et al. (2015)	-	-	✓	-	✓
Wang and Ding (2015)	-	✓	✓	✓	-
Nasir et al. (2015)	-	-	✓	✓	-
Day and Gunderson (2015)	-	-	✓	NS*	NS*
Hong et al. (2015a, 2015b)	-	✓	✓	✓	✓
Zhao et al. (2015)	-	-	✓	-	✓
Arora et al. (2015)	-	-	✓	✓	-
Labeodan et al. (2015, 2016)	-	-	✓	✓	-
Jain and Madamopoulos (2016)	-	-	-	-	-
Javed et al. (2016)	-	-	✓	✓	-
Mohammadmoradi et al. (2017)	-	-	✓	✓	-
Ekwevugbe et al. (2017)	-	-	✓	-	✓
Newsham et al. (2017)	-	-	✓	-	✓(Cubical)
Jin and Spanos (2017)	-	-	✓	✓	-
Wang et al. (2017)	-	-	✓	✓	-
Wang and Shao (2017a, 2017b)	-	✓	✓	✓	-
Nesa and Banerjee (2017)	-	-	✓	✓	-
Zhao et al. (2017)	-	-	✓	-	✓
Çiftler et al. (2017)	-	-	✓	✓	-
Wang et al. (2018)	-	-	✓	✓	-
Yang et al. (2018)	-	-	✓	-	✓(Cubical)
Mashuk et al. (2018)	-	-	✓	✓	-

*NS: Not specified in the paper

2.3.1 Occupancy Modeling using Statistical Methods

To apply statistical methods, large amount of data should be collected in different office buildings and over a long period to properly represent office occupancy. These methods analyze the collected data and the frequency of past events to fit probability distributions to parameters of interest (Ott and Longnecker, 2015). Having the distributions helps to estimate the probability of the occurrence of an action and create the occupancy model. Linear and logistic regression models, time series, and Bayesian estimates are examples of statistical methods.

According to Dodier et al. (2006), one of the biggest deficiencies in the determination of a reliable occupancy model is the lack of proper statistical analysis methods. Thus, they used a network of passive infrared occupancy sensors in two private offices and performed data analysis techniques based on Bayesian probability theory (a class of graphical probability models called belief networks) to determine the occupancy model. They showed that using probability models makes a significant improvement in the buildings' operation.

Using statistical methods, Chang and Hong (2013) defined the key parameters of the occupancy model as the average occupancy profile, the frequency of being absent from the office, and the absence duration. They collected the required data by installing 200 lighting-switch sensors for each cubicle office within open-plan offices on three floors of an office building. They found five typical occupancy patterns based on the differences in the daily occupant presence profiles. They claimed that the occupancy presence pattern is affected by the location of the cubicle in the office. Occupants in more isolated cubicles showed less movement. The same pattern was observed for the cubicles near the windows. The pattern with the highest number of occurrences in three-floor offices indicated that most of the occupants left their offices during the lunch break. Also, they found that the job category has a high impact on the occupancy pattern; however, due to privacy and security concerns they could not find more information for further investigation. Based on the gathered data, they generated uniform distributions of the number of daily absences of the occupants, the absence duration, and the start time of each absence. However, the start time of absence may not follow a uniform distribution. Although they tracked occupants in an open-plan office, they did not consider the effect of shared activities, such as meetings.

An object-oriented software module was introduced by Feng et al. (2015) using the occupancy models proposed by Page et al. (2008), Wang et al. (2011), and Chang and Hong (2013). The software includes all these models to have more comprehensive information to simulate different occupancy levels. They used the software for the simulation of a single-floor office building. The results were close to the input and the predetermined schedules for a typical day at the building level. However, there was a significant difference between the software derived from the occupancy module and that of the predetermined schedules at the room level. In addition, they concluded that occupant movements follow some statistical patterns, related to occupant job type or habits. They assumed the values of the inputs to the module; however, surveys and tracking techniques are required to provide reliable values for different types of inputs to the occupancy model.

2.3.2 Occupancy Modeling using Stochastic Methods

Occupancy models are highly dependent on the season, weather, time of day, and occupant habits and personality (Chang and Hong, 2013). Therefore, there is a significant need to consider the probabilistic modeling of occupant profiles to reflect these dependencies by leveraging various analysis methods (Haldi and Robinson, 2008). Stochastic models are developed by using real data related to occupant location, movement, and actions, collected over a short period. Stochastic analysis methods are then used to predict the probability of an event (i.e., occupant present in a space) to generate the stochastic profiles (Virote and Neves-Silva, 2012). Monte Carlo methods, Markov Chain, discrete and semi-hidden Markov Chain models, Poisson model, as well as state transition analysis are in this category.

Different occupant activities in office buildings are referred to as work states in this study. Determining the occupant's next work state based only on his/her present state is the basis of the Markov chain process. Yamaguchi et al. (2003) used probabilistic occupancy profiles within the development of a district energy system simulation model. They used the Markov chain to represent different work states and considered empirical distributions for the times of arrival, departure, and lunch break. In order to produce the Markov matrices which, define state transitions, two kinds of data were required: the duration of each work state and the distribution of work states. They assumed fix numbers for these two parameters; however, this information should be collected by conducting real-world experiments. This assumption results in simulating only one day and repeating it for the whole year without distinguishing between working days and weekends. Wang et al. (2005) proposed a probabilistic occupancy model in single-person offices. As in the research of Yamaguchi et al. (2003), they assumed that the duration of presence periods is time-independent (i.e., independent of the time of day). They found that the duration of intermediate absence periods follows an exponential distribution with one constant coefficient over a day. However, the occupied intervals are more complex and required two constant coefficients of the exponential distributions to simulate a sequence of alternating periods of absence and presence. They also considered that the times of the first arrival to the office, the last departure from the office, and the lunch break are normally distributed, which is not supported by their observations during the field experiments. Also, they treated all weekdays the same and did not consider long periods of absence, which leads to an overestimation of annual energy consumption.

To overcome the time-independence issue of previous occupancy models, Page et al. (2008) introduced probabilistic presence profiles as an input to a Markov chain to develop exponential distributions of intermediate periods of presence and absence with time-dependent coefficients. Their model also captures the changes in arrivals, departures, and typical breaks as well as periods of long absence. The only dependency of the proposed model to the occupant characteristics is related to occupancy inputs regarding the profile of probability of presence, parameter of mobility, and distribution of periods of long absence. Therefore, by providing correct and concise inputs to the model, it could be used for any building type with any occupant presence pattern. Beside its generality, the proposed model provides a more realistic estimation of the actual time spent by occupants in their zones and the number of their interactions with the environment. They estimated the occupancy pattern in a space if the presence of occupants is independent of each other. Also, the model eliminates the occurrence of undesired peaks that comes from repeating the same pattern for each occupant. However, the inputs to the model (e.g., the profiles of probability of presence and parameters of mobility) are very complex to obtain and define in simulation programs. Also, the model does not simulate the movement of occupants from one zone to another, which are of great importance to develop detailed occupancy models. To address this point, Tabak (2008) used a system of User Simulation of Space Utilization (USSU) to generate occupant activity and location in order to develop the movement patterns of the occupants in office buildings. However, the model was not capable of predicting the correct number of times that a workspace was occupied during a work day. Therefore, to improve the occupancy movement model, Wang et al. (2011) proposed a novel Markov chain approach to model stochastic occupancy of office buildings based on occupant movement among the spaces inside and outside a building. The model determines the location of each occupant and other key statistical properties of occupancy, such as the time of morning arrival and night departure, lunch time, periods of intermediate walking-around, etc. They claimed that the proposed occupancy model can realistically reproduce the occupancy distribution and the number of occupants. It also can be easily used to simulate occupancy for building energy simulation (especially HVAC system operation) due to its simplicity, accuracy and unrestraint nature. Although assuming Markovian property for the occupants' location and movement has been used for single-occupied offices, more validation is required especially for multi-occupied offices. They considered some assumptions in their modeling procedure to use the model for multi-occupied offices with no restrains related to the number of occupants and number of spaces within a building. These assumptions, in turn, lead to losing some inherent information about the occupants' movement. In addition, they could not calibrate and validate their model due to the lack of real measured data. Thus, they defined the inputs of the case study model based on experience.

A stochastic occupancy model was developed by Wang and Ding (2015) based on the correlation between the occupants' activities and equipment energy consumption. The accuracy of Markov chain models decreases when the amount of input data is increased. Thus, in order to alleviate this shortcoming, they used a combined model of Markov chain and the Monte Carlo methods (MCMC) to determine the occupants' activities and the computer input power for different time steps in multi-occupant office rooms. They used cameras to monitor occupant behavior and count their number each time step. The results showed a bimodal distribution for the average number of occupants over one week, which is compatible with the rules of building energy consumption. A building energy consumption prediction model can be generated using the occupant number. They also conducted a survey to gather data about occupants' working habits to determine different absence state. In the case of absent occupants, their equipment state (i.e., normal operation, standby, shutdown and locked) can be obtained based on their work habit. Occupant preferences regarding equipment usage patterns were gathered using power meters and the input power of equipment was recorded manually every 10 minutes. In addition, lighting and office equipment energy consumption was recorded based on the electricity consumption bills. They examined three office buildings with business, administration and scientific research functions. They reached a very low error rate (below 5%) between the predicted energy consumption from the model and actual energy consumption record. Despite of accurate representation of occupancy-based energy consumption prediction model, the proposed model is useful only for typical multi-occupied offices with more than eight occupants. In addition, meeting rooms, machine rooms, restaurants, exhibition rooms and other special function rooms are not included in their research. Counting the number of occupants and recognizing their activities are manual processes, which makes their tracking technology (cameras) inefficient for long-term tracking. In addition, the privacy concern regarding the usage of vision-based tracking systems restricts their implementation.

Dong and Lam (2011) developed a complex environmental sensor network to show the correlation between measured environmental conditions and occupancy status. Using a Gaussian Mixture Model (GMM) based on Hidden Markov Models (HMMs) resulted in detecting the number of occupants with 83% accuracy. The duration of occupancy was also calculated using a Semi Markov Model (SMM). To show the feasibility of the network, a case study was simulated producing occupancy data (i.e., the number of occupants and the duration of the occupancy). The results demonstrated 18.5% energy saving using perfect control for HVAC system. Although they tracked open-plan offices, the method only determines the number of occupants not their identities. In addition, their model is case-specific and only detects a maximum number of four occupants. The higher number of occupants results in lower accuracy and more complex computation.

2.3.3 Occupancy Modeling using Machin Learning Methods

Using approximate models guarantees the robustness of a system to the associated uncertainties. A robust system could react to uncertainties and accordingly tune itself. However, accurate models with low rates of prediction error are required to maximize the system's efficiency. Statistical methods, when used alone, cannot ensure robustness; however, both efficiency and robustness can be achieved when statistical methods are combined with approximate models. ML methods, also known as predictive analytics, use a combination of statistical and stochastic methods to analyze historical trends in the data, learn from them and then predict the future. There are various ML algorithms used in the context of Building Energy Performance (BEP) including decision tree, artificial neural networks (ANN), support vector machine (SVM), polynomial regression, and Bayesian networks (Tsanas and Xifara, 2012).

To leverage the statistical methods, different learning algorithms are integrated with historical trends to predict the future. Machine learning methods are getting increasing attention by scholars. Lam et al. (2009) and Dong et al. (2010) employed a complex sensor network to collect different parameters that are related to the occupancy presence in an open-plan office. They investigated the correlation between the detection of the number of occupants with those parameters to find the most important ones. Applying feature selection showed that CO_2 volume and acoustic level are the most important parameters in estimating the number of occupants. Therefore, they used these parameters as inputs to three occupancy estimation methods, namely SVM, ANN, and HMM. The results showed that the HMM is more accurate in terms of counting the number of occupants. Despite using an extensive network of sensors, the method did not show very robust performance in accurately estimating the number of occupants.

Yu (2010) used Genetic Programming (GP) to find the occupancy pattern in five single-occupied offices using motion sensors. GP is used to learn the occupants' behavior based on behavioral rules. They used the same sensor data (12 weeks of data) used by Page et al. (2008) to identify the state of the offices; however, they did not include weekends in their research. They considered different variables, such as time of the day and day of a week. They got these variables based on suggestions from Page et al. (2008) and Wang et al. (2005); however, they added two new variables to complement the learning procedure. These variables along with three random constants (e.g., day, hour, and minute) were combined using some operators to create behavioral rules. The prediction accuracy was selected as the fitness function. They trained several rules and found the best rule for each office. The best rule was then applied to the testing data. The results showed accuracy between 80-83%. This shows that the rules are robust, and GP is a proper algorithm for learning the occupants' behavior based on motion sensor data. The predicted probability of presence at the office follows the same trend of the recorded data, except for the final departure time, which results in overestimation of building energy consumption. They also found that the occupancy and vacancy intervals are exponentially distributed. They did not consider shared spaces in their research and the effect of proposed occupancy models on the operation of building systems.

Having a cubicle workstation equipped with different sensors (i.e., PIR motion sensor, CO_2, sound, light, and power use sensors), Hailemariam et al. (2011) used Decision Trees to investigate the relationship between different types of sensors. High accuracy as 97.9% was reached using motion sensor alone, which can be increased to 98.4% by considering multiple motion sensors. In contrast with many other research studies, they found that combining the data from different sensors worsened the accuracy of the occupancy detection. Localized occupancy detection in real-time for each cubicle workstation in an open-plan office was discussed in this research; however, there is no usage of this concept in the proposed method. The case study is limited to one cubicle and the effect of multiple occupants on the performance of the sensor network is neglected.

Ekwevugbe et al. (2017) used ANN for occupancy numbers estimation in multi-occupied offices. They used several sensors to gather indoor climate variables, energy data, and indoor events, such as PIR, CO_2, VOC, temperature, relative humidity (RH), acoustics, and light sensors, and camera. After processing the collected data, the feature selection process is performed to derive the most effective features from the sensor data, which is input to the occupancy profile estimation model. They showed that applying a sensor fusion process results in an optimized sensor selection and placement. Although they mentioned that the model has the potential to be linked to a control system, no further investigation has been performed to prove the performance of the model.

2.4 Control Systems

Lighting and HVAC systems and office equipment are the main sources of energy consumption in offices. Studies show that Americans and Europeans are spending on average 85% to 90% of their time in indoor environments (EPA, 1989; European Commission, 2003). In Canada, approximately 85% of the total energy in institutional and commercial buildings is consumed by heating, cooling, lighting, and IT equipment (NRCAN, 2012). Therefore, the intelligent use of energy within buildings is a recent trend of research studies and is the goal of Building Energy and Comfort Management (BECM) systems, which requires proper understanding of the interaction between occupants and building systems (Tabak, 2008; Yan, et al., 2015). The BECM system comprises HVAC system, lighting, hot water, and electricity control with the objective of fulfilling occupant requirements for comfort while reducing energy consumption during building operation (Tabak, 2008). In order to improve the building design and operation through BECM, proper energy conservation strategies should be considered. Applying control actions is an important part of the energy conservation strategies. Control actions include, but are not limited to, unplugging seldom-used appliances, enabling the "sleep mode" feature on computers, setting the thermostat to a reasonable temperature, using sunlight wisely, using blinds, etc. These control actions aim for a trade-off between minimizing the energy cost and usage while maximizing occupant comfort and satisfaction. However, current building control practices are unable to completely achieve these goals. This means applying more cost-efficient strategies can result in reducing the occupants' satisfaction and even productivity (Tabak, 2008; Singhvi et al., 2005).

2.4.1 HVAC Control Systems

Table 2-5 shows the comparison between different research studies applying HVAC control strategies, with the focus on occupancy tracking methods, occupancy modeling resolution, and occupants' preferences. Section 2.4.1.1 discusses the importance of utilizing occupancy tracking methods and occupancy information mentioned in Table 2-5 to control HVAC systems. Table 2-6 categorized the review papers based on the type of study and space, as well as control strategy level and setting. Three levels of control strategy resolution are considered in Table 2-6 including individual, zone, and room levels. A room refers to a space with four full-height walls, such as single- or multi-occupant offices and meeting rooms. A zone is part of a room and is defined according to either the number of HVAC terminal units or lighting fixtures in the room, unless otherwise is mentioned. For instance, a whole building or multiple rooms are defined as zones in some papers. Individual resolution is used whenever an individual control is available from HVAC and lighting points of view. For instance, an open plan office as a room could have multiple zones and there could be multiple individual sections (e.g., cubicles) within each zone. This classification is used throughout Section 2.4. Furthermore, to give a better insight regarding the HVAC control strategies, the type of control strategy and the resulting energy savings are provided in Table 2-7. Section 2.4.1.2 provides more detailed explanation regarding some of the references mentioned in Table 2-6 and Table 2-7 with emphasis on the method used to control the HVAC system (i.e., MPC). Section 2.4.1.3 discusses about the usage of simulation to provide a connection between occupancy models and Building Energy Management Systems (BEMSs) as well as to evaluate the energy performance of buildings due to the application of control strategies. Spatial resolution of the proposed HVAC control is discussed in Section 2.4.1.4 followed by the application of HVAC control strategies in real-world tests in Section 2.4.1.5.

2.4.1.1 Set point-based HVAC Control Using Occupancy Detection

In terms of controlling HVAC systems, occupancy related information is used for heat loads, system running time, required heating, cooling and distribution of conditioned air, and preferred temperature set points (Li et al., 2012). However, many current building control systems are designed based on regulations that assume maximum occupancy for all spaces at all times, regardless of the actual room occupancy. This results in unnecessary conditioning of spaces within a building, which ultimately leads to a large amount of energy losses. To alleviate this inefficiency and considering that HVAC systems consume about 50% of the total generated electricity in the U.S. (Erickson and Cerpa, 2010), smart control of HVAC systems has been proposed by many researchers. The control systems are set based on the knowledge regarding the occupants and their predicted usage patterns. Thus, a significant amount of energy could be saved using the control strategies (Erickson et al., 2009). In addition, HVAC systems are demand-driven operated. Occupied spaces should be ventilated in order to have proper air quality. Since ventilation depends on the number of occupants, the more occupants, the more ventilation is required. Some research papers showed that a reduction in the average ventilation rate in buildings that set ventilation rates based on maximum occupancy results in a decrease in the energy consumption by 10-15% while maintaining an acceptable indoor air quality (Erickson et al., 2009; Pavlovas, 2004; Feng et al., 2015).

The number of occupants, the occupancy duration, and the type of activity performed by the occupants are needed to calculate HVAC loads, system running time, required heating, and cooling and distribution of conditioned air. More personalized control strategies require occupant location and identity. This information is collected by means of monitoring technologies. Figure 2-2 shows the frequency of using different occupancy monitoring technologies by the 34 papers being reviewed with the focus on HVAC control strategies. As mentioned in Section 2.2, tracking technologies are used to derive occupancy information. Almost all the papers cited use of networks of different types of sensors to gather occupancy data. Half of all the papers mentioned the use of PIR, followed by CO_2, temperature and relative humidity sensors. Acoustic, lighting, and camera sensors had equal contributions. Pressure sensors and power meters were also utilized to improve occupancy detection (Dong and Lam 2011, 2014). Despite the high resolution available with the use of RFID tags, only 5% of research studies used this system as the tracking technology (Li et al., 2012). Therefore, among these papers, only a small number of them collected occupancy information with a high resolution. Most of the research studies focused only on the duration of the occupancy, which shows how long the room is occupied. In addition, a few studies determined the *x* and *y* coordinates of occupants as their location and only one paper used the occupants' identification for HVAC control application. Further, the number of studies used occupants' preferences to enhance the control strategy of the HVAC system is rather limited.

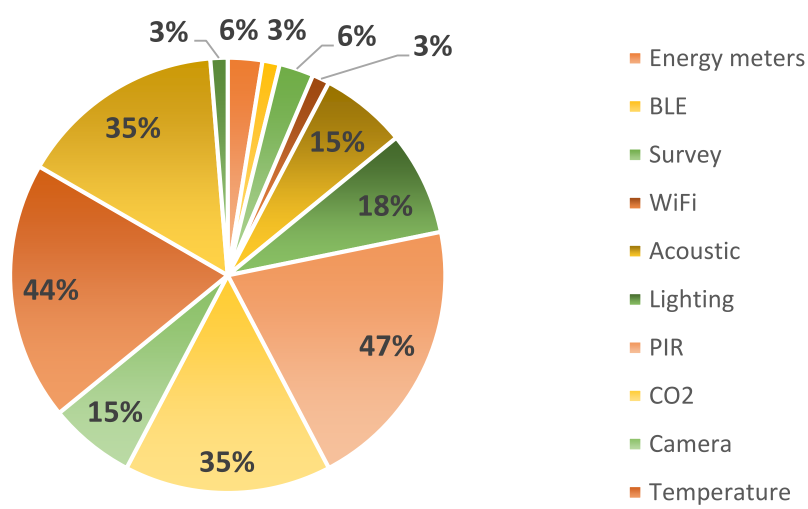


Figure 2-2 Percentage of the usage of different monitoring technologies in reviewed papers

Using monitoring technologies reveal the occupancy patterns, which show how the occupants use different spaces. Assigning groups of occupants with similar occupancy patterns to the same thermal zone was the basis of the start/stop operation of an HVAC system proposed by Capozzoli et al. (2017). They considered an office building with three thermal zones. Each thermal zone is composed of offices and corridor as the sub-zones. The typical occupancy profiles of each office sub-zone were found by means of ML techniques (i.e., K-means clustering algorithm and a binary decision tree (CART)). They also used optimization to find the optimal HVAC start/stop schedule. Simulation showed a 14% energy savings compared to an occupancy-independent operation schedule.

The same concept of matching occupancy patterns with thermal zone schedules was used by Wang et al. (2017). They tracked occupancy be means of a high-resolution occupancy detection, which works based on an iBeacon-enabled indoor positioning system. To avoid overcooling or insufficient cooling, they combined the occupancy profiles with a spatial dimension. This consideration provides the ability to reassign occupancy as a dynamic spatial occupancy distribution (DSOD) occupancy matrix. They used a feature-scaled artificial neural network algorithm to recognize the occupancy patterns from collected data. They compared the proposed control strategy with other traditional controllers by conducting a filed study and using computational fluid dynamics (CFD) simulation. Proper implementation of the proposed strategy showed 20% savings in energy consumption.

A more advanced control algorithm compared to the conventional on/off controllers was proposed by Foster et al. (2016). The control algorithm analyses the extracted data from a network of sensors, coordinates all the components of the system, and manages the communication between them. The sensor data determines the occupancy load of a room by counting the number of present occupants. A proper control signal is then sent to the HVAC system based on the sensor data. They used a microcontroller to implement the proposed control algorithm and reached almost 40% improved energy efficiency.

Another interesting point is pertinent to the evaluation of proposed control strategies. None of the reviewed papers applied cost-benefit analysis to justify the monetary benefits associated with the occupancy monitoring systems and the proposed control strategies. In addition, only two studies (i.e., West et al. (2014) and Brooks et al. (Brooks et al., 2014)) conducted surveys and utilized statistics methods to evaluate their control strategies according to the occupants' preferences and satisfaction level.

2.4.1.2 MPC of HVAC Systems Based on Occupancy Detection

Regarding the type of control strategy, set-point based methods using occupancy information and MPC were used by most of the researchers. Only a small number of the reviewed papers used optimization techniques and all of them applied optimization as a part of the MPC system. MPC is used in these studies to optimize the operation of the HVAC system (i.e., temperature set-points) by minimizing the energy consumption. Most of the studies considered the occupancy discomfort as a constraint in the optimization problem. This means MPC sets the space temperature in a way to minimize the total power consumption while restricting the hours of the occupants' discomfort. To do so, occupancy monitoring and modeling are used to determine the heat gain caused by the occupants as one of the main sources of the internal heat gains. For instance, a new predictive scheme for HVAC system was proposed by Majumdar et al. (2014) to make energy efficient decisions based on the past discomfort history of the occupants. Occupancy data were collected over the course of three months for a graduate office and laboratory, and for six months for a conference room using motion and CO_2 sensors. They assumed that the occupancy pattern would be similar for different weekdays but would differ between weekdays and weekends in offices and laboratories. To account for irregular occupancy of conference room, different occupancy profiles were used for different weekdays, including weekends. The predictive control strategy saved 7-10% of energy consumption while maintaining the occupancy comfort using simulation. Although they investigated the efficiency of the control system in shared spaces, the whole building was modeled as a single-zone, which restricts consideration of individual preferences regarding the temperature set point and local control.

In office buildings, where multiple rooms share a single variable air volume box, independent room conditioning (i.e., flow rate and temperature) is not possible. To control this type of HVAC system, which is called "*under-actuated*" system, two control algorithms were proposed by Brooks and Barooah (2014). The first algorithm is a modified version of the occupancy-based control algorithm proposed by Goyal et al. (2013) and the second one is an MPC algorithm based on the occupancy predictions. They used simulation to compare the proposed algorithms with the baseline algorithm and found 10-48% potential savings. Conducting experimental tests in two *under-actuated* zones, in which each of them has two rooms, showed that implementing the occupancy-based control algorithm results in 29-80% energy savings (Brooks et al., 2015). Despite the great energy saving potential of the proposed algorithms, their performance was not evaluated in open-plan offices, which makes the application of control strategies much harder compared to single-occupied offices. In addition, no tracking technologies were used to detect the real-time occupancy.

To improve the performance of standard linear MPC, a learning-based MPC technique was proposed by Aswani et al. (2012) to account for the impacts of occupancy by considering its fluctuations through learning. The proposed control strategy estimates occupancy using room temperature measured by temperature sensors. Implementing the control system in a single laboratory room showed that the proposed technique enhances the energy efficiency of the MPC while maintaining its robustness regarding the constraints satisfaction. However, the effect of control system on a HVAC system that serves multiple rooms as well as the application of local control in a shared space were not investigated.

2.4.1.3 Modeling HVAC Control Systems Using Simulation

Due to lack of a proper connection between occupancy models (i.e., occupancy patterns and preferences) and BEMSs, only few studies could achieve energy savings based on the probabilistic occupants' information. In order to overcome this problem, Dong and Andrews (2009) tried to provide this connection by using simulation tools. They developed a sensor-based network to model and predict occupant activities and connect them to BECM systems through simulation tools. By applying simulation and connecting the occupancy patterns (semi-Markov model) with HVAC system control (simply on/off the system), they obtained a 30% energy savings while maintaining a suitable indoor comfort level. Their method requires a large network of sensors to accurately detect occupant activities, and therefore, significant effort to code the events and then analyze them to find the actual occupancy pattern as well as pattern duration. All the parameters are defined empirically based on a predefined set of activities and any especial or unpredicted activity could not be captured. Thus, it can be used neither in prediction of occupancy to control building systems nor in other buildings and case studies. In addition, the complexity of the model increases with the size of monitored rooms, such as open-plan offices with many occupants and activities. This lowers the practically and tractability of their method for 'whole building' simulations.

Simulation was used in most studies to evaluate the performance of the control strategies. To consider the effect of temperature setback periods on building energy consumption, Gunay et al. (2016) developed an adaptive control strategy that learns the occupancy patterns and parameters, which describe the heat transfer process, to dynamically adjust the setback temperature schedules. They found that it takes less than two weeks for the control system to adapt to the occupancy patterns and temperature variations. Implementing the control strategy in a simulation model of a shared office space indicated 15-20% lower annual cooling loads and 8-10% lower annual heating loads.

2.4.1.4 Spatial Resolution and Local HVAC Control

Regarding the spatial resolution of the proposed control strategies, most of the reviewed papers evaluated the effect of HVAC control strategies on shared spaces. However, only one study investigated the effect of HVAC control strategy on an open-plan office (Dong and Lam, 2011). It can be seen that the spatial level of the control strategy is at zone level in most studies. However, nine of the 31 studies defined zones either as a room or multiple rooms, which lowers the accuracy of the application of the control strategy. Furthermore, the effect of individual preferences, which leads to implementation of local control, is not investigated.

Nagarathinam et al. (2017) investigated the spatial variations in temperature and occupancy on the HVAC system operation in open-plan offices. They used MPC to find the optimum temperature set point, which is later used in a proportional-integral-derivative (PID) controller to adjust the fan speed of the HVAC system with multiple AHUs. The aggregate occupancy count was determined through swipe-card meter. They assumed that the occupants are at their respective desks if present in the office space. Thus, the spatial location of each occupant was derived from the occupant's desk tagging information. Using simulation and comparing the proposed control strategy with static set-points based PID control strategies resulted in 12% energy savings. Although the effect of the occupants' movements on the energy consumption of the HVAC system is rather high, this effect is not considered in this research.

Applying local HVAC control strategies requires detailed occupancy information. The most important data is the specific location of occupants (i.e., the *x* and *y* coordinates), since the concept of local control is about considering the spatiotemporal variations of the space usage. Another important factor is related to the occupants' preferences. Thus, the identity of the occupants plays an important role in implementing the local HVAC control to identify their preferred temperature set points.

2.4.1.5 Application of HVAC Control Strategies in Real-world Tests

Using simulation by almost all the papers being reviewed shows the popularity and power of simulation tools to estimate building energy performance (Table 7). However, to bridge the gap between the simulation results and actual building energy consumption, many researchers used filed studies to investigate the effectiveness of their proposed control strategies. However, most of the field studies refer to the utilization of the monitoring technologies to gather occupancy data rather than applying the proposed control strategies. For instance, Agarwal et al. (2010) designed and implemented battery-operated wireless sensor nodes called Synergy Presence Nodes to accurately detect occupancy for individual offices. The sensors are low-cost, wireless, and easily deployable within existing buildings. They also have an estimated battery lifetime of over five years. They tested the proposed system by deploying it across ten offices over a period of two weeks. Using simulation showed significant energy saving potentials (i.e., from 10% to 15%) in HVAC system operation due to recognition vacancy periods. Despite practical aspects of their invented sensor network, since the system is attached to the offices' doors this system could accurately detect occupants only when they are near the doors. Thus, if the occupant is sitting at his/her desk there is going to be a delay in detection of actual room status. This results in some inconsistencies in predicted occupancy profile as compared to the actual one.

One instance of testing the system in the field can be seen in the work of Goyal et al. (2013) who compared the performance and complexity of three different control algorithms through simulations. The control algorithms improve both energy efficiency and thermal comfort of occupants using occupancy data. The first algorithm is an occupancy-based control algorithm that uses real-time occupancy measurements and zone temperatures to determine HVAC system set points and set back temperatures. The second and third algorithms work based on MPC and use occupancy data and predictions of occupancy, respectively. A baseline control algorithm commonly used in conventional HVAC systems was used to evaluate the performance of the proposed methods.

Table 2-5 Comparison of research papers applying HVAC control strategies with the focus on occupancy information (34 papers)

Reference	Occupancy Monitoring Method	Occupancy Model Resolution					Occupants' Preferences
		Location	Number	Identity	Duration	Activity	
Dong and Andrews (2009)	Acoustics, lighting, motion, CO_2, temperature and relative humidity sensors	-	-	-	✓	✓	-
Erickson et al. (2009)	Wireless camera sensor network	✓	✓	-	✓	-	-
Dong et al. (2011)	Acoustics, lighting, motion, CO_2, indoor and outdoor temperatures, relative humidity, wind speed sensors and pyranometer	-	✓	-	✓	-	-
Agarwal et al. (2010)	Synergy Presence Nodes (PIR and door sensors)	-	-	-	✓	-	-
Lo and Novoselac (2010)	-	-	-	-	-	-	-
Erickson and Cepra (2010)	Wireless camera sensor network	✓	✓	-	✓	-	-
Erickson et al. (2011)	SCOPES (wireless camera sensor network)	✓	✓	-	✓	-	-
Dong and Lam (2011)	CO_2, temperature, RH, acoustics, lighting, motion detection, pressure sensors and a network of cameras	-	✓	-	✓	-	-
Li et al. (2012)	RFID	✓	✓	✓	-	-	-
Aswani et al. (2012)	Room temperature sensor	-	-	-	✓	-	-
Goyal et al. (2012)	-	-	-	-	✓	-	-
Purdon et al. (2013)	PIR, temperature, and humidity sensors	-	-	-	✓	-	✓
Goyal et al. (2013)	-	-	-	-	✓	-	-
Balaji et al. (2013)	Wi-Fi and survey	✓	✓	✓	✓	-	✓
Oldewurtel et al. (2013)	Motion sensor	-	-	-	✓	-	-
Gunay et al. (2014)	-	-	-	-	✓	✓	-
Dobbs and Hencey (2014a, 2014b)	-	-	-	-	✓	-	-
	PIR motion detector network	-	-	-	✓	-	-
Majumdar et al. (2014)	Motion and CO_2 sensors	-	-	-	✓	-	✓
Bengea et al. (2014)	PIR, space temperature, humidity, CO_2, people counter, and supply temperature sensors	-	✓	-	✓	-	-
Gruber et al. (2014)	CO_2 sensors	-	✓	-	✓	-	-
Brooks et al. (2014)	PIR, temperature, humidity, CO_2 sensors, and web-based surveys	-	-	-	✓	-	✓
Dong and Lam (2014)	Temperature, RH, lighting, acoustics motion, CO_2 sensors, and power meters	✓	✓	-	✓	✓	-
West et al. (2014)	-	-	-	-	-	-	✓
Brooks and Barooah (2014)	-	-	-	-	-	-	-
Brooks et al. (2015)	PIR, temperature, humidity, and CO_2 sensors	-	-	-	✓	-	-
Goyal et al. (2015)	PIR, temperature, humidity, and CO_2 sensors	-	-	-	✓	-	-

Table 2-5 Comparison of research papers applying HVAC control strategies with the focus on occupancy information (34 papers) (Cont.)

Reference	Occupancy Monitoring Method	Occupancy Model Resolution					Occupants' Preferences
		Location	Number	Identity	Duration	Activity	
Foster et al. (2016)	Multiple sonic rangefinder modules (as motion sensor), smoke, acoustic, light, and temperature sensors	-	✓	-	✓	-	-
Gunay et al. (2016)	PIR, temperature, lighting sensors	-	-	-	✓	-	-
Lim et al. (2016)	-	✓	-	-	✓	✓	-
Capozzoli et al. (2017)	PIR, temperature, humidity sensors	-	✓	-	✓	-	-
Wang et al. (2017)	Cameras, temperature, humidity, CO_2 sensors, and BLE beacon	-	✓	-	✓	-	-
Nagarathinam et al. (2017)	Swipe-card meter	✓	✓	-	✓	-	-
Peng et al. (2017)	PIR, temperature, RH, and CO_2 sensors, energy meter	-	-	-	✓	-	-

Table 2-6 Comparison of research papers focusing on type of study and space, and the level of control strategy (34 papers)

Reference	Type of Study			Type of Space		Control Strategy Level		
	Optimization	Simulation	Field Study	Shared	Private	Individual	Zone	Room
Dong and Andrews (2009)	-	✓	-	✓ (Meeting room)	-	-	-	✓
Erickson et al. (2009)	-	✓	-	Hallways	-	-	✓	-
Dong et al. (2011)	-	✓	✓	✓	✓	-	✓	-
Agarwal et al. (2010)	-	✓	-	-	✓	-	-	✓
Lo and Novoselac (2010)	-	✓	-	✓	-	-	✓	-
Erickson and Cepra (2010)	-	✓	-	Hallways	-	-	✓	-
Erickson et al. (2011)	-	✓	-	Hallways	-	-	✓	-
Dong and Lam (2011)	-	✓	-	✓ (open-plan office)	-	-	✓	-
Li et al. (2012)	-	✓	-	✓	✓	-	✓	-
Aswani et al. (2012)	-	✓	✓	✓	-	-	-	✓
Goyal et al. (2012)	✓ (MPC)	✓	-	NS*	NS*	-	✓	-
Purdon et al. (2013)	-	✓	-	✓	-	-	✓ (multiple rooms)	-
Goyal et al. (2013)	-	✓	✓	-	✓	-	✓ zone=room	✓
Balaji et al. (2013)	-	-	✓	✓	✓	-	✓	-
Oldewurtel et al. (2013)	✓ (MPC)	✓	-	NS*	NS*	-	-	✓
Gunay et al. (2014)	-	✓	-	NS*	✓	-	-	✓
Dobbs and Hencey (2014a, 2014b)	✓ (MPC)	✓	-	NS*	NS*	-	✓ zone=building	-
Majumdar et al. (2014)	✓ (MPC)	✓	-	NS*	NS*	-	✓ zone=room	-
Bengea et al. (2014)	✓ (MPC)	-	✓	NS*	NS*	-	✓	-
Gruber et al. (2014)	-	✓	-	✓	-	-	-	✓
Brooks et al. (2014)	-	-	✓	✓	-	-	✓ zone=room	-
Dong and Lam (2014)	✓(NMPC)	✓	✓	✓	✓	-	✓ zone=room	-
West et al. (2014)	✓ (MPC)	-	✓	NS*	NS*	-	✓	-
Brooks and Barooah (2014)	✓ (MPC)	✓	-	-	✓	-	✓ (multiple rooms)	-
Brooks et al. (2015)	-	✓	✓	✓	-	-	✓ (multiple rooms)	-
Goyal et al. (2015)	-	-	✓	✓	-	-	✓ zone=room	✓
Foster et al. (2016)	-	✓	-	NA**	NA**	-	-	✓
Gunay et al. (2016)	-	✓	-	✓	-	-	-	✓
Lim et al. (2016)	✓	-	✓	✓	-	-	✓	-
Capozzoli et al. (2017)	✓	✓	-	✓	-	-	✓	-
Wang et al. (2017)	-	✓	-	✓(cubical)	-	-	✓	-
Nagarathinam et al. (2017)	✓ (MPC)	✓	-	✓	-	-	✓	-
Peng et al. (2017)	-	✓	-	✓	✓	-	✓	-

*NS: Not specified in the paper; **NA: Not applicable

Table 2-7 Comparison of research papers focusing on control strategy method and energy savings (34 papers)

Reference	Control Method	Energy Savings (%)
Dong and Andrews (2009)	on/off	30
Dong et al. (2011)	NMPC	18
Agarwal et al. (2010)	Set-point based control	10 - 15
Lo and Novoselac (2010)	CFD	12 or 30
Erickson et al. (2009)	Adaptive ventilation rate based on the number of occupants (demand-driven HVAC operation strategies)	14
Erickson and Cepra (2010)		20
Erickson et al. (2011)		42
Dong and Lam (2011)	Set-point based on occupancy schedule	19
Li et al. (2012)	Demand-Driven HVAC operation strategies	-
Aswani et al. (2012)	MPC	30-70
Goyal et al. (2012)	MPC	12-37
Purdon et al. (2013)	Set-point based control	60
Goyal et al. (2013)	MPC	42-60
Balaji et al. (2013)	Set-point based control	18
Oldewurtel et al. (2013)	MPC	34 and 50
Gunay et al. (2014)	Set-point based control	-
Dobbs and Hencey (2014a, 2014b)	MPC	37-44
		19
Majumdar et al. (2014)	MPC	7-10
Bengea et al. (2014)	Digital direct control and MPC	20-70
Gruber et al. (2014)	MPC and open-loop predictive controller	-
Brooks et al. (2014)	Set-point based control	37
Dong and Lam (2014)	NMPC	18 and 30
West et al. (2014)	MPC	19 and 32
Brooks and Barooah (2014)	Set-point based control and MPC	10-48
Brooks et al. (2015)	Set-point based control	29-80
Goyal et al. (2015)	Set-point based control and MPC	40
Foster et al. (2016)	Advanced Set-point based control	40
Gunay et al. (2016)	Dynamic setback temperature schedule	10 and 20
Lim et al. (2016)	Adaptive Temperature Control	12
Capozzoli et al. (2017)	Start/stop occupancy-based HVAC schedule	14
Wang et al. (2017)	Set-point based control	20
Nagarathinam et al. (2017)	MPC and PID	12
Peng et al. (2017)	Rule-based control	20

The baseline controller assumes occupancy during the day and vacancy during the night. While each room's air temperature is kept between lower and upper bounds, set back temperatures are used during nighttime. Comparing the proposed algorithms against the baseline algorithm showed significant energy savings with each of the proposed controllers. However, the feedback controller was found to be more suitable due to its simplicity and lower deployment cost compared to the more complex MPC-based controllers. A single room with two occupants was used to verify the simulation results of the first two proposed algorithms through a real-case experiment. The field study showed 40% energy savings. In addition, Brooks et al. (Brooks et al., 2014) implemented the occupancy-based control algorithm proposed by Goyal et al. (2013) in 12 shared rooms. The experiment demonstrated 37% energy reduction. Using web-based surveys before and after the implementation of the control algorithm showed no decreases in the occupants' comfort and air freshness (Brooks et al., 2014). Although multi-occupied rooms were used in these experimental studies, there was no accurate information pertinent to the actual number of present occupants. Assuming design occupancy (i.e., maximum number of occupants) for occupied rooms leads to overestimation of the building energy consumption.

2.4.2 Lighting Control Systems

Lighting systems consume about 20-45% of the total electricity consumption in office buildings (de Bakker et al., 2017) and are controlled by using occupancy sensors usually regardless of the occupants' activities. In this case, the sensors signal the state of the room (i.e., occupied or unoccupied) to turn on or off the lighting systems. Thus, the patterns of the lighting use are mainly related to the occupancy patterns in the office (Feng et al., 2015; Yun et al., 2012; Al-Mumin et al., 2003).

Applying lighting control systems helps maximize the energy efficiency of the lighting system. This can be done by using a set of presence sensors and actuators to control the operation of lighting system (e.g., turning on/off or dimming the lighting). An average energy saving of about 30% is claimed by applying lighting control strategies (Guo et al., 2010). In the following sections, research studies about different lighting control strategies are reviewed with regards to the type of monitoring systems, application of optimization methods, and local lighting control applications.

For instance, Van de Meugheuvel et al. (2014) used two different proportional-integral (PI) controllers to adjust dimming levels of multiple luminaires using the occupancy and lighting sensors. In one scenario, the PI controllers work independently based on the global occupancy information (classical PI controller). In the second scenario, the PI controllers communicate with each other and share information of the neighbor zones by adding networking capabilities to the controllers (i.e., networked PI controllers). They compared the power consumption of the proposed control strategies with optimum centralized control systems. The results showed that the energy consumption of the networked PI controllers is close to that of an optimized controller that operates based on the lighting inputs. The main drawback of this work is the lack of a field study to consider the effect of varying occupancy on the performance of the proposed controllers. As a result, they used static occupancy scenarios to evaluate the proposed method with pre-determined occupancy probabilities.

Table 2-8 shows an overview of the comparison between different studies applying lighting control strategies with the focus on monitoring methods, occupancy modeling resolution, and occupants' preferences. It can be seen from this table that almost all the studies tracked occupants to mainly determine the occupancy duration, which can be easily done by means of motion sensors. Hence, unlike papers that worked on the HVAC control strategies, most of the studies in this category only utilized motion and lighting sensors. There have been only two exceptions that used RFID and pressure, strain, and vibration sensors along with the motion sensors to have more detailed occupancy data. Furthermore, in recent studies there is a trend of using energy meters instead of occupancy sensors to track occupants for lighting control.

Table 2-9 categorizes the reviewed papers based on the type of study and space, and control strategy level and setting. Regarding the spatial resolutions, the lighting systems were controlled at individual level in most of the reviewed papers, which shows the emphasis on the application of local control. However, among all the reviewed papers, only five papers used surveys to infer specific occupant preferences. A large portion of the proposed control strategies (26 out of 37) have been implemented in real-life and the rest used simulation to investigate the performance of the control strategies.

In Table 2-10, the papers are classified based on the control strategy evaluation method and the obtained energy savings. One aspect that is not fully studied in these papers is related to the application of cost-benefit analysis. Only one study (i.e., Fernandes et al. (2014)) applied the cost-benefit analysis. Considering monitoring systems can be used for several purposes, such as security, facility management, safety and emergency situations, the cost-benefit analysis is required to investigate the balance between the cost of these systems and the gains of applying them in the real world.

2.4.2.1 Lighting Control Based on Binary Occupancy Detection and the Effect of Time Delay (TD)

As mentioned in Section 2.2.1, motion sensors are widely used to get binary occupancy data (i.e., whether an occupant is present in a specific space or not) and control the energy consuming systems in the building. Most of these sensors work based on the TD concept to control lighting systems. This means that whenever a motion is detected in a space under the coverage of the sensors the corresponding lighting system turns on. The lights will be switched "off" after a period of time has elapsed after the last motion is detected by the sensor. However, the TD is either pre-fixed or user adjustable to a fixed time. In addition, according to (Guo et al., 2010), there is an uncertainty associated with the occupancy data collection when single-point detection is used. Thus, long TD and high detector sensitivity settings are proposed to compensate for the uncertainty. However, a long TD results in energy usage during unoccupied periods and a short TD leads to occupant complains about false-negative errors (i.e., lights are switched off in occupied spaces due to the location of the occupant that is outside of the sensor field of view) (Guoet al., 2010).

Tiller et al. (2009) used three PIR sensors in 10 private offices as well as 23 cubicle workstations to collect data for 59 and 63 days, respectively. Comparing the occupancy profiles deduced from the collected data revealed considerable uncertainty in the measured data. As mentioned before, this uncertainty would result in using long TDs. Therefore, the effect of applying different TDs was investigated in private offices. The amount of energy that could be saved using 5- and 10-minute TD varied from 8.4 to 33.3% compared to 20-minute TD. Although big savings were achieved using the proposed method, its effectiveness is not evaluated in open-plan offices where lighting systems are shared between multiple occupants. They only claimed that the correlation between detecting occupants and the real occupancy in the open-plan office was weak when using PIR sensors.

Nagy et al. (2015) developed an adaptive lighting control system to determine occupant-specific set points for lighting system TD and illuminance thresholds. The control system adapts the TD and illuminance level based on the occupancy changes in each room. Implementation of the proposed control system in 10 different types of rooms in an office building shows that it took about one week for the control system to adapt to all occupants across all rooms. In addition, decreasing the TD by 5% resulted in doubling the energy savings without too much occupant discomfort. Although the results of their study showed the potential for increasing the energy savings by reducing the TD, their claim regarding the occupants' comfort is qualitative and based on not receiving complaints. However, a robust quantitative method is required to evaluate the occupants' comfort. They did not distinguish between different types of rooms, such as single-occupied, double-occupied, and multi-occupied offices. In addition, it is not clear how they defined the occupant-specific set points for offices with more than one occupant since different occupants have different preferences.

2.4.2.2 Lighting Control Based on More Advanced Occupancy Detection

Fixed presence sensors when used alone for controlling lighting system cause energy wastage due to ignoring the surrounding environment. Therefore, the energy efficiency of the lighting system was investigated by Delaney et al. (Delaney et al., 2009) using a WSN. To do so, they introduced LightWise (LIGHTing evaluation through WIreless SEnsor) to assess the lighting system of office buildings by determining points in which the energy wastage occurred. A light detector and PIR sensor were used in their study to detect ambient light and luminaries' state (i.e., lights being on or off) and the occupants' presence, respectively. Testing took place in three separate spaces, a large open plan office, a small individual office and a corridor. They proposed two control strategies, presence detecting and manual switch control strategies. They found that 50-70% energy saving can be achieved by either replacing this system with traditional fixed presence sensors or optimizing the current system based on the obtained results regarding the potential points where of the energy consumption can be improved in office buildings.

On the basis of work done by Harris and Cahill (2005) mentioned in Section 2.2.2, Harle and Hopper (2008) used CAPM system to control lighting system. They employed an ultrasonic location system with 95% accuracy to detect occupancy in 36 offices, 6 corridors, and 9 communal rooms. They collected data for a year and used the data from 60 working days randomly selected from a year for evaluation purposes. To identify the tracked occupants, the room outliers, ingress and egress zones were defined as spatial zones. Three scenarios were applied to measure lighting energy consumption: (1) keep lights on 24 hours a day; (2) turn on the lights after the first arrival of the office owner and off after the last person who leaves the room; and (3) automatically turn on/off lights using location-aware system. Comparing the results of the last two lighting schemes showed a 50% saving in energy consumption. The main problem with the proposed system is that the lighting system would switch on whenever a person enters a room regardless of the size of the room. This provides unnecessary lighting in the case of spacious rooms.

Manzoor et al. (2012) combined passive RFID technology with PIR sensors to provide more accurate occupancy detection in open-plan offices. The proposed approach resulted in more energy efficient lighting control. Although they reached 13% energy savings, they used RFID gateway attached to the office entrance door to show the number of occupants who enter or leave the office. Thus, there is no indication of occupants' behavior within the office and their preferences on the lighting control. Another problem with the proposed method is regarding its implementation that requires the installation of RFID reader.

2.4.2.3 Daylight Harvesting

Another strategy to save more energy is through the control of lighting systems while wisely utilizing the ambient natural light present in a space. This strategy is called daylight harvesting. Using this strategy leads to energy savings by dimming or switching off the lighting whenever sufficient ambient light is present (Si et al., 2017).

Galasiu et al. (2007) studied the potential savings by applying three different lighting control strategies simultaneously and independently in an open-plan office over a period of one year. Using occupancy sensors, external light sources (i.e., daylight harvesting), and individual dimming controls for each occupant independently resulted in 35%, 20%, and 11% energy savings, respectively. However, 42-47% energy savings can be achieved by combining these three strategies compared to using the same lighting system without controls. In addition, comparison between the applications of three control strategies with the energy usage of a conventional lighting system, in which the lighting is always on during the working hours, showed a 67-69% reduction in energy consumption. They concluded that in the case of using only one control strategy, the occupancy sensor would be the best choice. Although different control options were considered in this study, the effect of the lighting energy savings on the thermal performance of the office was not investigated. Since there is a relationship between the internal heat gains and the lighting energy consumption, the effect of applying lighting control strategies on the performance of HVAC system should be considered to reach even greater energy savings.

Yun et al. (2012) presented the results of a survey conducted in four offices to monitor occupancy patterns, lighting system usage, and lighting system energy consumption. No statistically significant relationships were found between external illuminance and lighting use patterns. However, some clear effects related to the time of day were observed, such as a strong tendency of turning on lighting when occupants arrive in the morning. They found 43% reduction in the lighting energy consumption when using automatic dimming control. On the other hand, there was up to 50% increase in lighting energy use when considering a change in occupancy patterns.

Zhu et al. (2017) proposed a simulation methodology that uses energy meters to derive the occupant schedules. They evaluated lighting control strategies to determine the potential energy savings based on four different occupant profiles. Energy savings of almost 62% were achieved as a result of switching from conventional lighting systems to lights with daylight-responsive dimming functions.

Kuo et al. (2017) designed an automated lighting control system that adjust the indoor illuminance level using the individual preferences, natural light, and shading system. They implemented the proposed system in a scaled physical model and tried to adjust the indoor illuminance level to reach a pre-selected target value. The implementation showed that the control system is only able to reach the desired light level in the case of bright outdoor conditions. The use of real occupant data (i.e., presence and preference regarding indoor illuminance level) and thermal control of the indoor conditioning are mentioned as areas for future investigation.

2.4.2.4 Optimization of lighting systems

Wen and Agogino (2008) developed an intelligent lighting optimization algorithm to implement lighting control with the objectives of providing both energy efficiency and occupant satisfaction. They implemented the proposed framework in a shared office using a wireless networked lighting system. They reached 68% and 48% energy savings for a sparsely (four occupants) and a more densely (seven occupants) occupied office, respectively. There is no indication of the working hours and occupancy pattern in the two scenarios used to evaluate the proposed framework. More reliable data could be collected using occupancy tracking and lighting sensors. They considered daylight harvesting, occupancy control, and lighting level tuning in another optimization problem. The optimization algorithm generates light output for each light fixture based on occupant requirements while minimizing the energy consumption. 60% energy saving was found by implementing the proposed lighting system in a small open-plan office (Wen and Agogino, 2011). The main shortcoming of the proposed model is the lack of occupancy sensors to detect the real occupancy pattern in the open-plan office. In addition, lighting preferences are assumed and predetermined instead of asking the occupants to provide their requirements.

Rossi et al. (2015) proposed an optimization framework to determine the dimming levels of multiple lighting fixtures in an open-plan office under two control scenarios. In the first scenario, the target illumination levels are predefined as 500 lux and 300 lux for occupied and unoccupied zones, respectively. In the second scenario, the target illuminance levels are determined based on occupant desires. They used three approaches called minimum, maximum, and average approaches to calculate the target illuminance levels in the case of multiple occupants with different desired illuminance levels. They tested the performance of the control scenarios by simulating the same open-plan office considered in Van de Meugheuvel et al. (2014). The first control scenario showed no overshoot/undershoot with a small settling time to reach to the final steady-state value. They found that the first control scenario results in almost the same energy savings compared to a benchmark model. Although they proposed different approaches in the second scenario, they did not use the actual occupancy of the open-plan office. Thus, the effect of different occupancy patterns is not investigated.

Caicedo and Pandharipande (2016) optimized the lighting power consumption of an open-plan office using a central controller system. They used the same model of Caicedo et al. (2015) (i.e., a hypothetical open-plan office with 24 zones) to compare the performance of the dual-beam luminaires with a standard-beam lighting system. Each zone has one occupant and is equipped with zone-level luminaires, and lighting and occupancy sensors. Two optimization scenarios were defined: one with illumination and dimming levels constraints and the other with only illumination constraints. The comparison of the two optimization scenarios showed better spatial uniformity of the dimming level for the first scenario in both absence and presence of daylight. They also measured the target illumination level of each lighting fixture during the calibration procedure and compared the illumination level achieved by each optimization scenario with the target level. The results demonstrated that both scenarios were able to reach the target illumination level for all luminaires. They did not implement the proposed method in a real open-plan office and only used a simulation model to investigate the feasibility and effect of the method on the energy consumption savings.

Caicedo et al. (2017) proposed a lighting control algorithm that determines the dimming level of luminaires based on the collected sensing data to achieve a desired illumination level in each workspace (zone) of an open-plan office. They applied the proposed model in a real open-plan office with eight luminaires associated with eight occupancy sensors, eight ceiling light sensors, and eight workspace wireless lighting sensors. They compared the achieved illumination levels of the eight workspaces using a control method that receives inputs from ceiling and workspace lighting sensors (combined control) with those of control methods that either receives inputs from ceiling lighting sensors or workspace lighting sensors. They concluded that the combined control saves more energy and produces more robust results than the case of using ceiling lighting sensor data. In both of these research studies, there is neither an indication of the model architecture nor the software/tools that were used to implement the proposed model. Although they mentioned that each zone is tracked by the occupancy sensor, they did not mention the type of sensor. In addition, the occupancy data were not collected, and they assumed an occupancy scenario with only four zones being occupied for both the simulation model and the testbed measurements.

2.4.2.5 Local Lighting Control

Although significant energy savings can be achieved by applying lighting control strategies, the different occupant preferences regarding lighting and visual comfort are usually overlooked, or even compromised, especially for open-plan office buildings. Therefore, a localized lighting control algorithm was proposed by Labeodan et al. (2015, 2016) using occupancy data from pressure sensors, and its performance was compared to that of the lighting control based on dual-PIR sensor data. The pressure sensors collected more accurate and reliable occupancy data. However, one important limitation of this type of sensor is their inability to detect occupancy when the occupants are walking or standing in the room. Although open-plan offices were considered, the effect of neighboring occupants' presence and their preferences was not considered in this study.

The effect of using task lights in reducing the energy consumption and improving the occupants' satisfaction was investigated by Lim et al. (2017). They used illumination loggers to track the lighting usage of two office spaces, one with the daylight and the other one with only the artificial lighting. After placing the task lights, a visual comfort survey was conducted to evaluate the occupant lighting preferences. Comparing the energy saving potential in two offices showed 78% lighting energy saving in the case of using daylight. The main limitation of this study is the lack of occupancy sensors to collect the real offices' occupancy data.

2.5 Roadmap for Cognitive Building Management

One of the long-term goals of the building industry is to design and operate cognitive buildings in a way that could satisfy occupants' comfort requirements, enhance the performance of energy-consuming systems, and increase efficiency. To reach these goals, there is a need for: (1) comprehensive information pertinent to different building systems; (2) real-time data collection; (3) proper management of the collected data (i.e., cleansing, storing, and mining); and (4) data-driven decision models to act upon the collected data and modelled information for integration and coordination of different building systems (Pasini et al., 2016). Having this holistic framework provides better insight regarding the current and future states of buildings and their evolution towards more intelligent and responsive entities. In order to achieve this goal, research and development should be integrated with technological advances.

Table 2-8 Comparison of research papers applying lighting control strategies with the focus on occupancy monitoring method and occupant preferences (37 papers)

Reference	Monitoring method		Occupancy Model Resolution				Occupant Preferences
	Occupancy Monitoring	Lighting Sensor	Location	Number	Identity	Activity	
Garg and Bansal (2000)	Smart TD sensor	-	-	-	-	-	-
Jennings et al. (2000)	Ultrasonic and PIR sensors	-	-	-	-	-	-
Escuyer and Fontoynont (2001)	Motion sensor	✓	-	-	-	-	✓
Maniccia et al. (2001)	PIR	Photosensor	-	-	-	-	-
Von Neida et al. (2001)	PIR	Photosensor	-	-	-	-	-
Chung and Burnett (2001)	Motion sensor and observation	Photosensor	-	-	-	-	-
Jennings et al. (2002)	Ultrasonic and PIR sensors	-	-	-	-	-	-
Galasiu et al. (2007)	Motion sensor	Photosensor	-	-	-	-	-
Wen and Agogino (2011)	-	-	-	-	-	-	✓
Harle and Hopper (2008)	Ultrasonic sensors	-	✓	-	-	-	-
Mahdavi et al. (2008)	Motion, light, temperature, and RH sensors and photography	-	-	-	-	✓	-
Galasiu and Newsham (2009)	Motion sensor	✓	-	-	-	-	-
Delaney et al. (2009)	PIR	✓	-	-	-	-	-
Tiller et al. (2009)	PIR	-	-	-	-	-	-
Rubinstein and Enscoe (2010)	Motion sensor	-	-	-	-	-	-
Pandharipande and Caicedo (2011)	Ultrasonic sensors	Photosensor	-	-	-	-	-
Wen and Agogino (2011)	-	-	-	-	-	-	-
Manzoor et al. (2012)	PIR, RFID	-	✓	✓	✓	-	-
Oldewurtel et al. (2013)	Motion sensor	-	-	-	-	-	-
Fernandes et al. (2014)	Infrared and ultrasonic sensors	Photosensor	-	-	-	✓	-
Aghemo et al. (2014)	PIR	photosensor	-	-	-	-	-
Van de Meugheuvel et al. (2014)	Motion sensor	✓	-	-	-	-	-
Peruffo et al. (2015)	Motion sensor	✓	-	-	-	-	-
Rossi et al. (2015)	Motion sensor	✓	-	-	-	-	✓
Nagy et al. (2015)	PIR	✓	-	-	-	✓	-
Caicedo et al. (2015)	Motion sensor	✓	-	-	-	-	-
Pandharipande and Caicedo (2015)	Motion sensor	✓	-	-	-	-	-
Caicedo and Pandharipande (2016)	Motion sensor	✓	-	-	-	-	-
Nagy et al. (2016)	PIR	✓	-	-	-	-	✓
Labeodan et al. (2015, 2016)	Pressure, strain, vibration, and PIR	-	-	✓	-	-	-
Caicedo et al. (2017)	Motion sensor	✓	-	-	-	-	-
Lim et al. (2017)	-	-	-	-	-	-	✓
Zhu et al. (2017)	Energy meters	-	-	✓	-	-	-
Delgoshaei et al. (2017)	Energy meters	-	-	-	-	-	-
Gentile and Dubois (2017)	-	-	-	-	-	-	-
Dikel et al. (2017)	Motion sensor and pressure mat	✓	✓	-	-	-	-

Table 2-9 Comparison of research papers applying lighting control strategies with the focus on type of study and space, control strategy level and setting (37 papers)

Reference	Type of Study			Type of Space		Control Strategy Level			Control Strategy Setting	
	Optimization	**Simulation**	**Field Study**	**Shared**	**Private**	**Individual**	**Zone**	**Room**	**TD**	**Illuminance Setting**
Garg and Bansal (2000)	-	-	✓	-	✓	-	-	✓	✓	-
Jennings et al. (2000)	-	-	✓	-	✓	-	✓	-	-	✓
Escuyer and Fontoynont (2001)	-	-	✓	NS	NS	NS	NS	NS	-	✓
Maniccia et al. (2001)	-	✓	✓	✓	✓	-	-	✓	✓	-
Von Neida et al. (2001)	-	✓	✓	✓	✓	-	-	✓	✓	-
Chung and Burnett (2001)	-	✓	✓	✓	-	-	✓	-	✓	-
Jennings et al. (2002)	-	-	✓ (cubical)	✓	-	-	✓	-	✓	-
Galasiu et al. (2007)	-	-	✓ (cubical)	✓	-	✓	-	-	-	✓
Wen and Agogino (2011)	✓	✓	✓ (cubical)	✓	-	✓	-	-	-	✓
Harle and Hopper (2008)	-	-	✓	✓	-	-	✓	-	-	-
Mahdavi et al. (2008)	-	✓	✓	✓	✓	-	-	✓	-	-
Galasiu and Newsham (2009)	-	-	✓ (cubical)	✓	-	✓	-	-	-	✓
Delaney et al. (2009)	-	-	✓	✓	✓	-	✓	-	-	-
Tiller et al. (2009)	-	-	✓ (cubical)	✓	✓	✓	-	✓	-	-
Rubinstein and Enscoe (2010)	-	-	✓ (cubical)	✓	-	✓	-	-	-	✓
Pandharipande and Caicedo (2011)	✓	✓	-	✓	-	✓	-	-	-	✓
Wen and Agogino (2011)	✓	-	✓ (cubical)	✓	-	✓	-	-	-	✓
Manzoor et al. (2012)	-	-	✓ (cubical)	✓	-	✓	-	-	✓	-
Oldewurtel et al. (2013)	✓	✓	-	NS	NS	-	✓	-	-	-
Fernandes et al. (2014)	-	-	✓	✓	-	-	-	✓	✓	✓
Aghemo et al. (2014)	-	-	✓	NS	NS	-	-	✓	✓	✓
Van de Meugheuvel et al. (2014)	✓	✓	-	-	-	-	-	✓	-	✓
Peruffo et al. (2015)	-	✓	-	-	-	✓	-	-	-	✓
Rossi et al. (2015)	✓	✓	-	✓	-	✓	-	-	-	✓
Nagy et al. (2015)	-	-	✓	✓	✓	-	-	✓	✓	✓
Caicedo et al. (2015)	-	✓	-	✓	-	✓	-	-	-	✓
Pandharipande and Caicedo (2015)	✓	✓	-	✓	-	✓	-	-	-	✓
Caicedo and Pandharipande (2016)	✓	✓	-	✓	-	✓	-	-	-	✓
Nagy et al. (2016)	-	-	✓	✓	✓	-	-	✓	✓	-
Labeodan et al. (2015, 2016)	-	-	✓(Laboratory)	✓	-	✓	-	-	✓	-
Caicedo et al. (2017)	-	-	✓ (Laboratory)	✓	-	✓	-	-	-	✓
Lim et al. (2017)	-	-	✓	✓	-	✓	-	-	-	-
Zhu et al. (2017)	-	✓	-	✓	-	-	✓	-	-	✓
Delgoshaei et al. (2017)	-	-	✓	✓	-	-	-	✓	-	-
Gentile and Dubois (2017)	-	✓	-	-	✓	-	-	✓	-	✓
Dikel et al. (2017)	-	-	✓	✓	-	-	✓	-	✓	✓

Table 2-10 Comparison of research papers applying lighting control strategies with the focus on control strategy evaluation and energy savings (37 papers)

Reference	Control Strategy Evaluation Occupant Feedback		Energy Savings (%)
	Method of Collecting Feedback	Statistics	
Garg and Bansal (2000)	-	-	5
Jennings et al. (2000)	-	-	20-26
Escuyer and Fontoynont (2001)	Survey-Interview	✓	-
Maniccia et al. (2001)	-	✓	17-60 (irregular occupied spaces) 28-38 (private offices)
Von Neida et al. (2001)	-	✓	17-60 (irregular occupied spaces) 28-38 (private offices)
Chung and Burnett (2001)	Observation	-	26-39 (lights on for 14 hours) 6-23 (manual control)
Jennings et al. (2002)	-	-	10-20
Galasiu et al. (2007)	-	-	up to 69
Wen and Agogino (2011)	-	-	up to 68
Harle and Hopper (2008)	-	-	50
Mahdavi et al. (2008)	-	✓	66-71
Galasiu and Newsham (2009)	Survey	-	up to 32
Delaney et al. (2009)	-	-	50-70
Tiller et al. (2009)	-	-	8-33
Rubinstein and Enscoe (2010)	Survey	✓	40
Pandharipande and Caicedo (2011)	-	-	-
Wen and Agogino (2011)	-	-	60
Manzoor et al. (2012)	-	-	-
Oldewurtel et al. (2013)	-	-	up to 34
Fernandes et al. (2014)	-	-	28-33
Aghemo et al. (2014)	Questionnaire	✓	17-32
Van de Meugheuvel et al. (2014)	-	-	-
Peruffo et al. (2015)	-	-	-
Rossi et al. (2015)	-	✓	20-45
Nagy et al. (2015)	-	✓	23-38
Caicedo et al. (2015)	-	-	-
Pandharipande and Caicedo (2015)	-	-	10-40
Caicedo and Pandharipande (2016)	-	-	23-54
Nagy et al. (2016)	Questionnaire	✓	13
Labeodan et al. (2015, 2016)	-	-	-
Caicedo et al. (2017)	-	-	-
Lim et al. (2017)	Survey	✓	78
Zhu et al. (2017)	-	-	62
Delgoshaei et al. (2017)	-	✓	23
Gentile and Dubois (2017)	-	✓	30-55
Dikel et al. (2017)	-	-	79

2.5.1 Building Management Evolution

A BMS is defined as a control system consisting of software, hardware and communication protocols to monitor and control a vast range of building systems (Papantoniou et al., 2015). Traditionally, building systems are operated separately. Each system is monitored and controlled regardless of the conditions of other building systems, and the different types of data collected from different sources are not shared. However, the increase in the number of systems and technological advances have led to the development of integrated BMSs for automated building management, where different building systems are connected to one another via a centralized management system (Doukas et al., 2007). This means that the BMS allows automating the building systems adjustments (Somayajulu, 2014). In spite of the power of BMSs in automatically controlling building systems, there are two major problems when using them. Firstly, these systems require human input to function, such as selecting the right time to turn on the lights. Secondly, the BMSs are very complex in terms of operation. Therefore, considering the vast range of parameters affecting the energy performance of buildings, achieving an optimal operation (i.e. minimum energy cost and maximum comfort levels of the occupants) using BMS is difficult. To address these needs, more energy-efficient systems and new technologies are required in buildings to identify the sources of energy waste and occupant discomfort and react accordingly to individual, organizational and environmental requirements. One promising solution that can achieve these goals is the integration of IoT with BMS, which enables *smart or intelligent buildings* (Wong et al., 2005; Roselli et al., 2015). The main capability of the IoT paradigm is integrating sensing, communication, computation, and control (Patel et al., 2016). In this paradigm, each system has its own computing component, which can communicate and interact with other systems through either *cloud computing* or *edge computing* (Lilis et al., 2017). Computing at the edge of an IoT architecture is one of the most recent types of sensor data processing. Edge processing can help overcome latency and other issues that come from using centralized cloud computing (Zhao et al., 2018). Hence, the IoT-equipped systems can host sensors and actuators and can be controlled based on distributed decision-making. For instance, a smart building equipped with IoT can detect an increase in the occupancy rate of space, and accordingly adjust the building systems. These types of buildings are also called *context-aware buildings* that could decide when to make the necessary adjustment to different building systems by considering all parameters affecting the performance of the building. In addition, the growing integration of Artificial Intelligence (AI) predictive analytics with smart BMSs makes building systems self-learning and intelligent in terms of adapting to changes within the building. Integrating IoT with AI and cognitive learning would result in *CBM*, which is autonomously aware of the energy performance of the building and its occupants' comfort level. This type of BMSs learns from building systems' operation patterns and the occupants' behaviors to optimize energy performance and improves the occupants' satisfaction. Therefore, cognitive buildings have three main capabilities: (1) having information regarding the building performance and its components' conditions (e.g., occupants' comfort levels) through the application of advanced data analytics to near real-time data gathered by IoT sensors; (2) learning building operational patterns along with the occupants' requirements and preferences and recognizing any unexpected changes; (3) deploying changes to building systems' settings considering occupants comfort levels. New levels of productivity, increasing environmental efficiency, enabling new business models and improving occupant well-being are some of the advantages of shifting to CBM (Somayajulu, 2014; IBM, 2016).

Another main gap in the application of the current BMSs is the lack of proper communication and data exchange between different systems. For instance, the gathered data from occupancy monitoring technologies, which can be used for energy management, are not shared with other building systems, such as security and emergency management systems nor are they saved for further analysis. Through the application of IoT, the collected data from different resources can be shared and used for various purposes. The Crystal building in Singapore (The Crystal, 2016), the Edge building in Amsterdam (Bloomberg, 2015), the Capital tower in Singapore (CapotaLand, 2017), the Al Bahar towers in Abu Dhabi (Al Bustani, 2014), and the Well Living Lab in U.S. (IBM, 2017) are examples of buildings using the IoT in different BMSs.

2.5.2 Proposed Roadmap

Based on the above discussion, a roadmap towards CBM (IBM, 2017) is proposed in this section. The proposed roadmap shows the evolution paths towards the CBM vision by integrating different research areas with advances in Information and Communication Technology (ICT). In this roadmap, three main steps are required for the realization of this vision: (1) *Technologies*: Adopt, deploy and integrate emerging technologies, such as IoT-based sensor networks; (2) *Methods and Analytics*: Extract the required information and patterns from the collected data using different techniques to add higher level of intelligence to BMSs; and (3) *Goals*: Define the gaps in the BMSs and the goals to fill these gaps for achieving CBM.

The proposed roadmap comprises four branches showing the goals of CBM including near real-time sensor information, ontological Occupant Information Modeling (OIM), dynamic occupancy prediction, and adaptive operation systems. The full realization of CBM requires achieving all these goals as shown in Figure 2-3. The overall view of the paths toward CBM and the areas that require further development are illustrated in this figure. It is important to mention that the proposed roadmap fits with the previous sections. Section 2 provided a review pertinent to different occupancy monitoring and sensing techniques. Research review of occupancy modeling and control of operating systems are covered in Sections 3 and 4, respectively. Each of the roadmap branches is explained in the following paragraphs.

- **Near Real-time Sensor Information**

As discussed in Section 2, different sensing technologies are utilized to monitor environmental and occupancy parameters affecting the energy performance of buildings. The IoT network can provide seamless sensing and control by: (1) continuously collecting the necessary data in near real-time, (2) processing and analyzing the sensor data while benefiting from the information in the OIM for occupancy prediction. The results of this analysis provide the input for the adaptive operation systems, and (3) autonomously communicating the results to actuators for controlling different building systems through IoT-based BMSs. The near real-time sensor information will result in improving the efficiency and cost-effectiveness of the building (Santucci, 2010). For instance, the energy consumption of a building can be optimized through the application of near real-time local control strategies.

- **Ontological OIM**

As discussed in Section 3, different models for predicting occupancy and occupant behavior in office buildings are developed to quantify the impact of occupant-related parameters on building energy consumption. However, the lack of standardization and consistency in these models makes it difficult to compare them with each other. To address this problem, IEA EBC Annex 66 *Definition and Simulation of Occupant Behavior in Buildings* was created to investigate the shortcomings of occupancy models and find the inconsistencies in them (IEA EBC Annex 66, 2013-2017). For instance, Hong et al. (2015a, 2015b) focused on energy-related building occupant behavior and suggested an ontology called *Drivers-Needs-Actions-Systems (DNAS)* framework to standardize the energy-related occupant behavior modeling. The proposed ontology is based on *need-action-event* cognitive theoretical frameworks that are presented over the past 40 years to represent the interactions of occupants with building systems. The occupancy models try to capture the stochastic nature of occupant behaviors by providing a connection between the occupant "*inside world*" inputs (drivers and physical, physiological or psychological needs) and the environmental "*outside world*" outputs (actions and events). To represent the proposed *DNAS* framework in an interoperable language, an Extensible Markup Language (XML) schema named *occupant behavior XML* (obXML) is used to capture the data syntax and structure and present them in a standardized way. Using this schema provides an interface to integrate the *DNAS* framework with the building energy simulation tools.

On the other hand, Building Information Modeling (BIM) is a shared digital representation of a building and its functional objects. BIM basically hosts a database of information embedded within spatial objects. BIM has an open standard called Industry Foundation Classes (IFC) (Pasini et al., 2016). This open-BIM has a standard representation of all types of buildings components and their properties, and it can support the interoperability between different BMSs (buildingSMART, 2018).

As an extension of the abovementioned occupant ontology, and in order to accommodate and share the great amount of sensor data that will be collected in the CBM systems of the future, it is important to represent the IoT devices (i.e. sensors and actuators) and the collected sensor data in BIM. The OIM should be developed based on a detailed study of occupancy ontology (i.e., occupancy features and the relationships between them). Eventually, the new entities and relationships of the OIM can be represented in IFC as part of the open-BIM (Hong et al., 2015b; Energy Information Administration (EIA), 2010). This fusing of occupant-related information into open BIM will contribute to the CBM by facilitating the interoperability of different BMSs.

- **Dynamic Occupancy Prediction**

As explained in Section 3, occupancy prediction models are developed using the data collected by occupancy sensors during the occupancy monitoring period. The most advanced occupancy prediction methods use ANN techniques to capture the hidden patterns in the collected data using iterations. These techniques are assumption-independent, which makes their predictive power very strong and reliable (Srivastava, 2015; Lejlic, 2017). Deep Learning (DL) (Bengio, 2009) is a new type of ANN that can structure algorithms in layers and learn on its own without the need for the manual steps of extracting relevant features of the input data (Amato et al., 2017). Instead, the input data are directly fed into the DL model, which extracts the most discriminative features and combinations of features (Amato et al., 2017; Pingel, 2017). DL techniques use the back-propagation algorithm to discover intricate structures in large data sets. They determine how the internal parameters of a model should change to compute the representation in each layer from the representation in the previous layer and perform predictions at the near-human level of accuracy (Amato et al., 2017; LeCun et al., 2015). Therefore, these new techniques can be employed to develop the next generation of occupancy models, which can predict the behavior of occupants with a high level of accuracy. The resulting predictive behavior modeling along with the information from sensors and the OIM eventually result in dynamic occupancy prediction as one of the CBM goals.

- **Adaptive Operation Systems**

The main advantage of integrating IoT, BIM, and OIM in BMS is the application of adaptive operation systems. For example, local control strategies can contribute to energy conservation by combining the spatiotemporal variations of space usage with occupant information. The integration of the research related to the above three branches (i.e., sensor information, OIM for occupancy prediction) eventually leads to consistent and continuous assessment of building performance by providing real-time information pertinent to the conditions of the building and its occupancy (Pasini et al., 2016). In this case, IoT-based self-tuned systems collect the information from the sensors and use context-aware analytics to achieve distributed decision making, which sends proper control signals to building systems to locally perform the control actions (Ersue et al., 2015). In this way, the building learns from the collected data and occupant information including comfort preferences and fine-tune its systems for optimal efficiency.

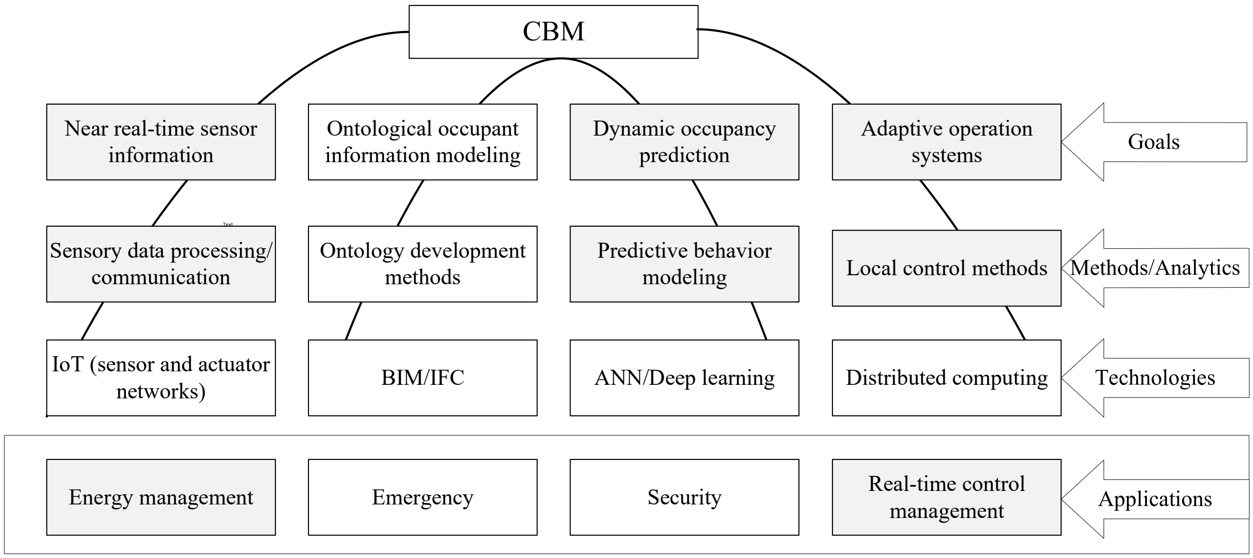


Figure 2-3 Roadmap for Cognitive Building Management

2.6 Summary and Conclusions

This chapter provided a comprehensive critical review that covers all the dimensions explained in Section 1.2 with respect to office buildings' energy management. Since the application of occupancy detection systems and occupancy-based control strategies differ based on the nature of the building (e.g., residential vs. commercial buildings), the focus of this chapter was only on office buildings, especially open-plan offices. The added value of the chapter relies on its comprehensiveness and linkage between different dimensions of the research. In addition, a roadmap regarding the advances in different dimensions was presented. The proposed roadmap provides a high-level view of the directions for future research towards CBM. By integrating all the components in the roadmap, a vision of CBM can be seen where buildings' systems, their occupants, and all other stakeholders have intelligent support from systems encapsulating sensor data and control strategies. The benefits of a CBM are: (1) the integration of the IoT with BIM and BEMS to change buildings from adaptive and predictive to cognitive and energy-efficient entities; (2) real-time monitoring of the energy consumption and occupants' behavior to reduce energy consumption; and (3) the integration of sensor networks and cloud-based technologies in the built environment and their future applications, such as safety, emergency, and security applications.

CHAPTER 3 RESEARCH FRAMEWORK

3.1 Introduction

As discussed in Section 1.2, the application of different intelligent local control strategies on building energy-consuming systems will result in reducing the building's energy consumption and improving occupancy comfort. In order to have effective control strategies, the most important factors affecting the operation of building systems should be investigated. Considering the fact that occupants spend over 80% of their time within indoor environments makes occupancy a paramount parameter in the evaluation of building energy consumption (Zhu et al., 2005). Moreover, according to (Yu et al., 2011), among the factors influencing the total building energy consumption, building occupants' presence and preferences could have high impacts on the energy usage of a building. Focusing on office buildings, most of the energy is consumed during working hours (Masoso and Grobler, 2010). Lights are often set to produce more light than necessary and HVAC systems are set based on the peak occupancy regardless of actual space utilization pattern. This makes the occupants-related parameters driving factors causing large discrepancies in the building energy usage even between similar buildings with the same function and located at similar locations. Therefore, these factors should be considered as accurate as possible when dealing with the building operation and energy models. As a result, the research in this area is two-folded. Firstly, due to the vital impact of occupancy data on building operation, accurate occupancy prediction models should be developed. A good occupancy prediction model requires enough amount of input data pertinent to the occupants' presence and preferences, which show the space utilization patterns and desired settings of the building systems, respectively. Secondly, control strategies should be generated based on the occupants' presence and preferences.

In addition, generally, there is an inverse relationship between the energy consumption of operational systems and the comfort level of occupants using these systems. Occupants' preferences regarding the energy-consuming systems affect their energy consumption. On the other hand, changing the settings of these systems has an impact on how occupants feel about their surrounding conditions. As a result, finding a balance between these two important concepts is crucial to improve the building operation. Optimal operation of building energy-consuming systems is a complex procedure for decision-makers, especially in terms of minimizing the energy cost and the occupants' discomfort. Proper control strategies should be selected to optimally operate energy-consuming systems while minimizing their energy usage and the occupants' discomfort considering different constraints. Simulation techniques can be used to investigate the effect of different control strategies on building energy consumption and the occupants' satisfaction. This is done by performing sensitivity analysis on the settings of the energy-consuming systems to find how changes in the settings of these systems affect the performance of the simulation model. However, simulation alone cannot explore the whole search space of a complex energy efficiency problem; therefore, optimization methods are required to fully investigate all the possible different combinations of settings.

The application of the near real-time local control strategies improves the building performance by providing varying control actions depending on the dynamic occupancy information on the performance of the buildings' operation systems. To this end, the occupants' detailed information is collected using a new monitoring technology (i.e., Bluetooth RTLS). After developing the personal profile for each occupant using advanced data analysis, a BEMS applies the near real-time local control of HVAC and lighting systems. The BEMS needs to solve an optimization problem with the input of the occupancy model (i.e., occupants' profiles and preferences) and energy efficiency requirements in building codes and standards (e.g., the HVAC system setback temperature, minimum lighting level, etc.). The results of the optimization are the settings of the HVAC and lighting systems that will provide minimum energy consumption and maximum occupants' satisfaction. The integration of the simulation model and the optimization algorithm allows exploiting the best features of these tools simultaneously.

The overall proposed framework of this study is discussed in this chapter, which includes introducing the two main modules comprising the proposed framework.

3.2 Research Methodology

3.2.1 Overview of the Research Framework

In order to achieve the objectives of minimizing the building energy consumption as well as the occupants' discomfort hours, a detailed building energy simulation model should be developed and encapsulated within an optimization algorithm. Since the operation of building systems are highly dependent on the presence of occupants, the integrated model should select the most optimized settings for building systems based on the dynamic space occupancy information. Having reliable insight regarding the occupancy information is especially crucial when applying local control strategies in shared spaces. When applying local control of building systems, the modeled space should be divided into multiple zones to assign relevant dynamic occupancy information to each zone. The zoning is applied to consider the effect of (1) different types of activities performed in each zone; (2) different number of the HVAC terminal units as will be discussed in the following section; (3) different facade orientation for perimeter zones. In large open-plan offices occupied with multi occupants, space should be divided into zones for the adoption of the local control strategies (Salimi et al., 2017). The concept of the proposed local control strategy is demonstrated in Figure 3-1. The figure shows a multi-occupied open-plan office, which is equipped with RTLS, which tracks occupants and captures their location, and activities at the zone level and at a certain frequency over time. To control the HVAC and lighting systems at a more detailed level, the office can be divided into different zones according to the number of HVAC terminal units or the number of lights. Knowing the location of a specific occupant, the corresponding HVAC terminal unit and corresponding light are adjusted using local control strategies.

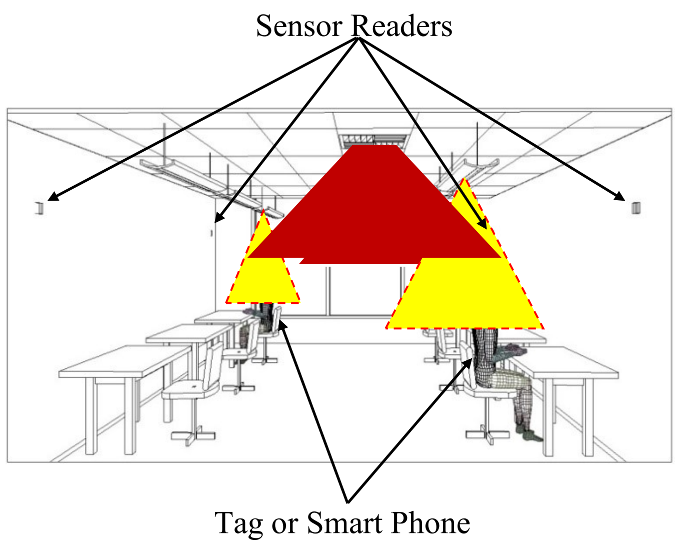


Figure 3-1 Local control strategy with occupancy monitoring (Liu et al., 2016)

In this study, the proposed methodology comprises two main modules: (1) simulation-based multi-objective optimization module and (2) occupancy module as illustrated in Figure 3-2. The processes required for each module are shown in this figure. Firstly, information regarding space occupancy is obtained through the occupancy module. To do so, real occupancy data should be collected over a reasonable period using RTLSs within the occupancy module. The collected data are then processed and the derived information is imported to the simulation model as an indication of the real occupancy space utilization patterns. This information helps the model to better differentiate occupants in the monitored shared space.

The next module, which is simulation-based multi-objective optimization, starts with developing a detailed simulation model to evaluate the energy performance of the building using the building and its energy-consuming systems' characteristics. Feeding the energy simulation model with the occupancy information, a simulation-based optimization problem is then solved to determine the values of the decision variables, which are the settings of the building energy-consuming systems. These values are calculated based on the problem objective functions. The output of this module is a file containing information pertinent to the local control of the building systems.

3.2.2 Occupancy Module

There are many factors determining the accuracy of the occupancy model including the occupants' identities, the duration of the occupants' presence, their locations in different zones of a building, and their preferences. New RTLSs can provide the location and duration of presence while the preference data can be collected by a simple survey. The occupancy module is used to determine the occupants-specific dynamic profiles based on their presence data as shown in Figure 3-2. The main benefits of having dynamic occupancy profiles, which reveals the occupant's information regarding his/her location and space utilization pattern are: (1) unlike models that rely on averaging the various occupants' behaviors or schedules, the dynamic occupancy profiles can capture the diversity of the different occupants' behaviors, which is very important factor in open-plan offices; (2) real-time monitoring and the resulting decision making are the closest ways to emulate the real behavior of occupants and their interaction with building's energy-consuming systems. The dynamic occupancy profiles can distinguish between different occupants' schedules and habits. These profiles can be used to effectively apply occupants' personalized preferences. More details regarding the development of the occupancy module are presented in Chapter 4. Moreover, the performance of the proposed occupancy model is investigated in Chapter 5 using different sensitivity analyses.

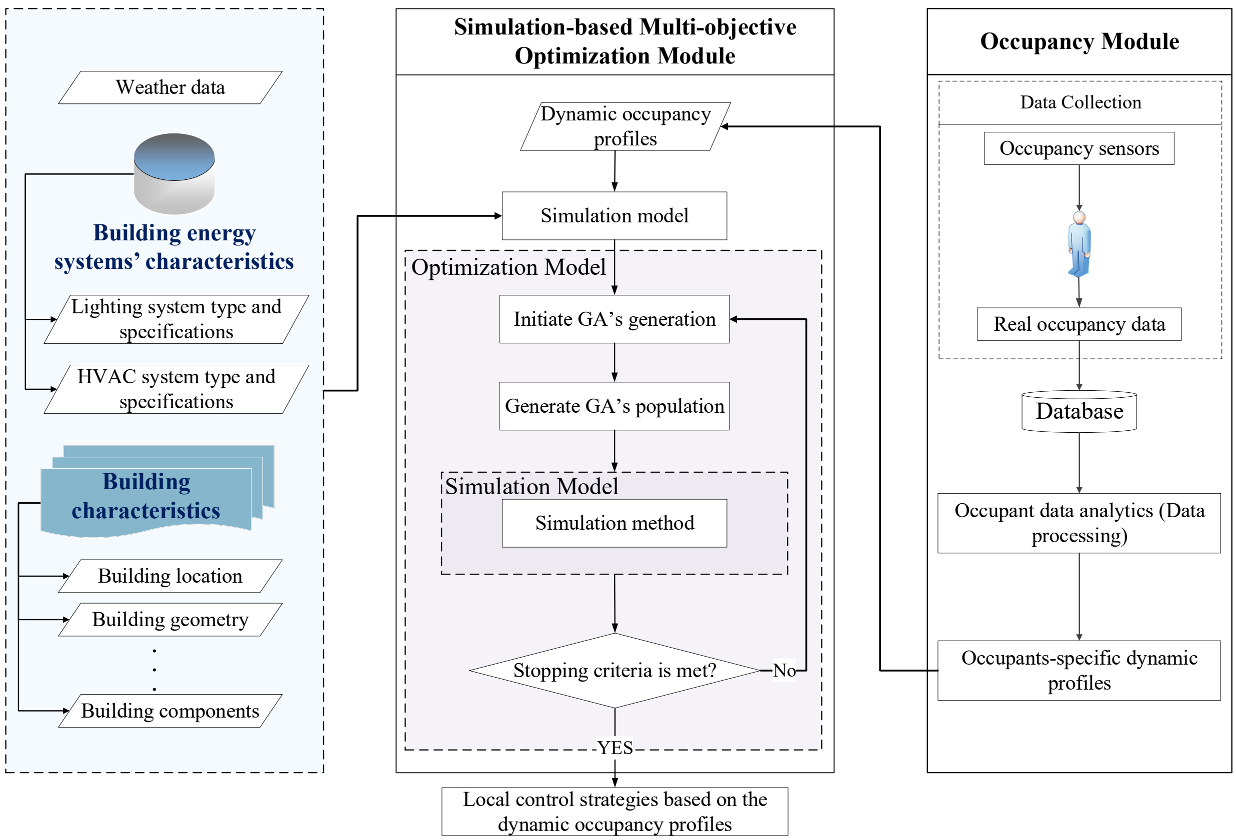


Figure 3-2 Proposed framework main modules

3.2.3 Simulation-based Multi-objective Optimization Module

Due to different parameters affecting buildings' energy consumption as discussed in Section 1.1, the optimal operation of buildings' energy-consuming systems is a complex procedure for decision-makers especially in terms of minimizing the energy cost and the occupants' discomfort. They have to find proper control strategies to optimally operate energy-consuming systems while minimizing their energy usage and the occupants' discomfort considering different constraints. Generally, there is an inverse relationship between the energy consumption of operational systems and the comfort level of occupants using these systems. Occupants' preferences regarding energy consuming systems affect their energy consumption. On the other hand, changing the settings of these systems has an impact on how occupants feel about their surrounding conditions. As a result, finding a balance between these two important concepts is crucial to improving the building operation. Simulation techniques can be used to investigate the effect of different control strategies on the building's energy consumption and the occupants' satisfaction. This is done by performing sensitivity analysis on the settings of the energy-consuming systems to find how changes in the settings of these systems affect the performance of the simulation model. However, simulation alone cannot explore the whole search space of a complex energy efficiency problem; therefore, optimization methods are required to fully investigate all the possible different combinations of settings. To this aim, Chapter 6 provides an in-depth discussion of the integration process and elaborates on the algorithms for the application of occupancy-centered local control strategies.

3.3 Summary

Addressing the current research gaps regarding the efficient and intelligent energy management of buildings, an overview of the proposed framework has been discussed in this chapter. The methodology, which consists of several modules and phases, covers the development of a new adaptive probabilistic occupancy model. This is done using prediction techniques. Moreover, the dynamic occupancy profiles, derived from the RTLS data, are fed to the simulation-based optimization model to assess the effect of different intelligent and occupancy-centered local control strategies on the building's energy-consuming systems and the occupants' satisfaction. In the upcoming chapters, each module of the proposed method is explained in detail and validated using case studies.

CHAPTER 4 PROBABILISTIC OCCUPANCY PREDICTION MODEL

4.1 Introduction

As discussed in Section 1.1, among the parameters affecting the energy consumption in buildings the ones that vary with time, such as the occupancy, play an important role in accurately evaluating the energy performance of buildings. However, these parameters are difficult to predict due to the uncertainties associated with them.

Occupancy models, which are derived based on space utilization patterns and occupant behavior, are key factors to accurately estimate the energy consumption of buildings. According to (Feng et al., 2015; Yan et al., 2015), there are different resolution levels for occupancy models, which are highly context-dependent. These levels should be determined according to the required granularity of occupancy models used for different purposes. For instance, a finer level of granularity is needed to apply lighting control strategies. Given that HVAC systems need some time to adjust the indoor temperature to a specified target set-point, less accuracy in occupancy detection may not lead to a significant thermal discomfort (Shen et al., 2017). The high-resolution occupancy models provide the following information: (1) the location of occupants, (2) their identities, (3) the number of occupants in each zone of the building, and (4) their activities at each time-step. Having this information helps to determine the occupants' interactions with building systems (Hong et al., 2015a). This will eventually lead to the application of occupancy-centered local control strategies on the systems. Furthermore, occupancy-related information is useful for different energy/comfort management purposes as well as other areas, such as safety/security, space management, and emergency responses.

Based on the above discussion, occupancy modeling is a complicated procedure and many occupant behavior analytics (data processing) steps are required to polish the input data and create a reliable occupancy model. Monitoring and data collection are important steps to develop a detailed occupancy model. A good occupancy model requires enough input data pertinent to the occupants' space utilization patterns. This data is gathered for a reasonable period through monitoring techniques, such as different RTLSs. However, most of the occupancy detection systems cannot provide the number of occupants and the specific location of each occupant (i.e., the x and y coordinates of the occupant) when they are used for open-plan offices. Most of the research works that consider shared multi-occupied offices did not distinguish between different individuals. Therefore, their practicality is reduced for open-plan offices, which have multiple thermal zones (Li et al., 2012). In addition, they lack detailed investigation of the effect of the individual preferences of occupants sharing the same area on the energy consumption of the building. Therefore, there is a need to use proper sensing and occupancy modeling techniques to distinguish between different occupants in multi-occupied offices and apply their preferences.

This chapter aims to develop a new adaptive probabilistic occupancy prediction model for open-plan offices based on occupancy data. In this study, the occupancy modeling (i.e., occupants' profiles) has been further enhanced using an inhomogeneous Markov chain prediction model, which distinguishes the temporal behavior of different occupants within an open-plan office based on occupancy space utilization patterns data. To this end, the occupants' detailed data (who, where, when) is collected using a relatively new monitoring technology (i.e., Bluetooth RTLS) that responds to occupancy changes in open-plan office buildings with acceptable accuracy. After developing the personal profile for each occupant with varying time-steps using advanced data analytics, a new adaptive probabilistic occupancy prediction model is developed to be used for occupancy prediction of open-plan offices. The proposed model is verified using a case study. Finally, comparing the building's real occupancy and produced results by the occupancy prediction model provides the validation of the applicability of the proposed model.

4.2 Research Methodology for Developing Occupancy Model

There are many factors determining the accuracy of the occupancy model including the occupants' identities, the duration of the occupants' presence, their locations in different zones of a building, and their preferences. New RTLSs can provide the identity, location, and duration of presence while the preference data can be collected by a survey. The zoning concept plays an important role in capturing the detailed occupancy information in real open-plan offices and improving the accuracy of the occupancy prediction model. Open-plan offices should be divided into multiple zones to assign different probabilistic occupancy information to each zone. The zoning is applied to consider the effect of (1) different types of activities performed in each zone; (2) different number of the HVAC terminal units or the number of luminaires; (3) different facade orientation for perimeter zones, to name a few. The occupancy prediction model is used to determine the occupants-specific probabilistic profiles based on their presence data. The main benefits of this dynamic and probabilistic occupancy prediction model are: (1) Probabilistic feature benefit: unlike models that rely on averaging the various occupants' schedules, the probabilistic occupancy prediction model can capture the diversity of the different occupants' presence patterns using stochastic methods, which is very important factor in open-plan offices; (2) Dynamic update benefit: real-time monitoring and the resulting decision making are the closest way to emulate the real patterns of occupants' presence and their interaction with building systems. The probabilistic occupancy prediction model can distinguish between different occupants' schedules and habits.

To consider the variations in the occupants' profiles due to their temporal behavior, each day is divided into different time slots. There are typical events of importance in office buildings that should be captured while defining these time slots, such as the first arrival to the office. These time slots are determined based on the patterns seen in the collected data as will be explained in Section 4.4.3. The events of importance indicate the typical patterns of the occupants' activities in open-plan offices. These activities are referred to as work states in this study as shown in Table 4-1. The duration of each work state is determined using the monitoring data. The first arrival to the office is defined as the first reading of the occupant's presence in the office after his/her long absence during the night. The last departure from the office is determined as the point when there is no recording of the occupants' presence for a duration greater than four hours after that point. Lunch break is defined as a break happening around noon with a duration greater than half an hour. Other breaks during the day with duration shorter than half an hour are considered as short breaks. Meetings, as one example of long breaks, are events that are happening based on a predefined schedule, such as weekly, bi-weekly, etc.

Table 4-1 Typical Occupancy Work States in Office Buildings

Work State	Description	Label (in Prediction Model)
1	Working in Occupant's Station (Occupant's Zone)	S_{oc}
2	Working in Other Occupants' Station (Other Zones)	S_{ot}
3	Lunch Break (lb)	S_{lb}
4	Short Break (sb)	S_{sb}
5	Long Break/Meeting (lm)	S_{lm}

Figure 4-1 shows the proposed framework of developing a new adaptive probabilistic occupancy prediction model. This framework comprises three main steps including data collection, data processing, and occupancy prediction model. Data collection is discussed in the following section. During the occupant behavior analytics (data processing), the analysis is required to find important occupancy features, such as the number of present occupants, periods of absence and presence, and other occasional variations in the occupants' profiles. Calculation of the occupancy rate is explained in Section 4.2.2.1 after discussion regarding the data processing phases in Section 4.2.2. The development procedure of the occupancy prediction model is then explained in Section 4.3.

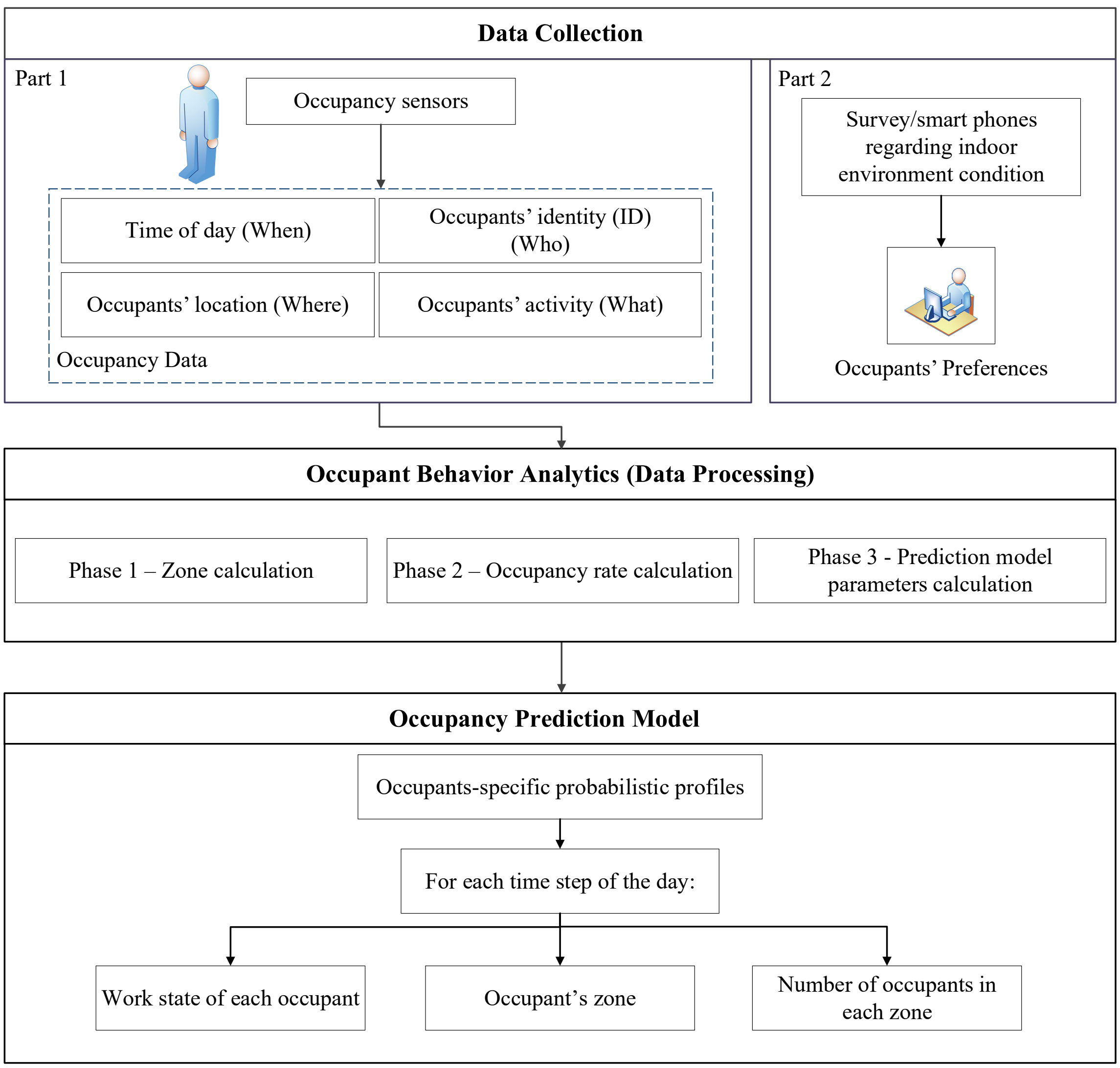


Figure 4-1 Occupancy module components

4.2.1 Data Collection

Considering that the energy performance of an open-plan office is mainly influenced by its occupants, the office occupants' schedules and habits determine the input to the occupancy prediction model. To this end, the office occupants need to be monitored. The data collection step includes two parts to gather all important occupancy data. As mentioned in Section 4.1, occupants should be monitored over a reasonable period using RTLS to get occupants' locations, their identities, presence time, and the type of activities and find the spatiotemporal patterns of the occupant's behavior (part 1). RTLSs are wireless systems that are used to automatically identify and monitor the location of objects or people in a defined space at a point in time that is or is close to real-time (Curran et al., 2011). RTLS comprises of different components: (1) various tags and badges or cell phones to send signals to the sensors (locators); (2) locators for reading tags; (3) platforms (Infrared, Ultrasound, Radio Frequency, and others); (4) timing cables or wireless bridges, for the connectivity of sensors with each other and with the host computer; (5) location engine, for calculating tag's position using various techniques; and (6) end-user software application for recording data (Akanmu et al., 2013).

To monitor occupants within an open-plan office using Bluetooth technology, either tags or cell phones can be used. These tags will send signals to the RTLS. Each tag has a unique ID number; thus, when a person moves in the area covered by locators, the system detects the unique ID number of the tag and measures the direction of a radio signal transmitted by the tag. Using Angle-of-Arrival (AoA) signal processing method, the incidence angles of the received signals are calculated with respect to the known positions of the locators. Applying a triangulation method, the position of tags can be determined (Azzouzi et al., 2011). As illustrated in Figure 4-2, an accurate 3D position is determined using at least two locators. In practical applications, several locators should be used, depending on the size of the monitored office, to detect the tags providing continuous positioning and substantially improving the accuracy and reliability of the results (Quuppa, 2017). The main targets of the data collection procedure are the office occupants who are assigned to the office. Thus, visitors, who may enter the office during the day, do not interfere with the data collection procedure and would not affect the accuracy or performance of the proposed model since there is no tag associated with them. The proposed occupancy prediction model is developed based on the office occupants' data and will predict the future occupancy profiles of the office occupants. In addition, occupants are questioned regarding the settings of the building energy-consuming systems to know their preferences (part 2).

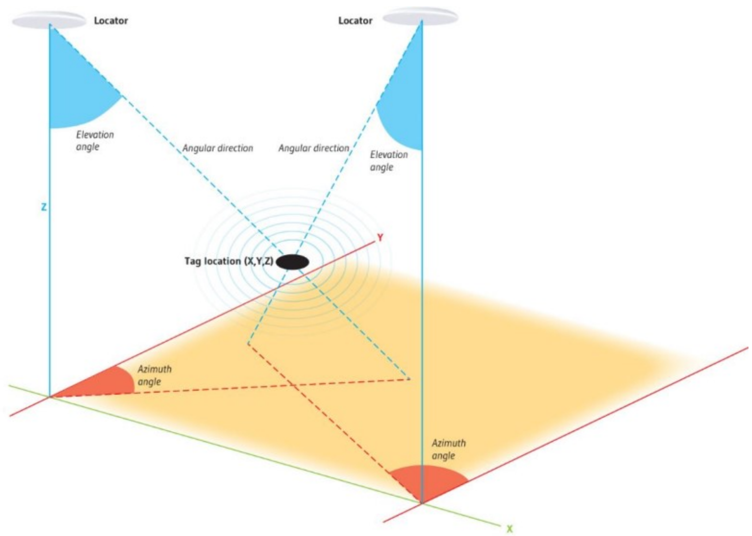


Figure 4-2 3D Positioning using RTLS (Quuppa, 2017)

4.2.2 Occupancy Behavior Analytics (Data Processing)

The occupancy behavior analytics (data processing), which is comprised of three phases, is performed to find important occupancy features, such as the number of present occupants, periods of absence and presence, and other occasional variations in the occupants' profiles. Figure 4-3 depicts the pseudocode showing how the collected data is converted to the occupancy location and presence duration information to calculate the occupancy zone and rate, respectively. All the phases of the data processing procedure are also shown in Figure 4-4.

The occupant's zones for each time-step of the total daily presence time (PT) are determined at the end of Phase 1 according to the x and y coordinates of his/her tag for each time-step. Using the information from phase one, the number of present occupants, and eventually the occupancy rate of the office, as will be explained in Section 4.2.2.1, are determined at zone and room levels for each time-step of the total occupancy duration (TOD), each day of a week (d), and for the total number of weeks of the data collection (W) during phase 2. Phase 3 of the data processing focuses on the analysis required to obtain the parameters that are necessary for developing the occupant-specific transition probability matrices, as will be explained in Section 4.3.1, for each time-step of each day of a week. This phase starts with changing the time-step resolution according to the purpose of the occupancy model. For instance, HVAC system local control strategies require longer time-steps knowing that it takes time for the system to adjust the zone temperature.

```
Set W = the total number of weeks of the data collection, w = week, D = the total days of a week, d = the day
of a week, o = occupant, N_o = the total number of occupants in the office, t^dr = detection time resolution, z =
occupant zone, n_o = the number of present occupants, occ_r = occupancy rate

For each d in D
  For each w in W
    For each o in N_o
        T^o_start = the first time that the occupant o is detected in the morning
        T^o_end = the last time that the occupant o is detected in the evening
        PT^o = T^o_end - T^o_start
        For t_i^dr in PT^o
          if there are multiple readings for each t_i^dr:
              calculate the average coordinates of readings with the same t_i^dr
          if there is a missing data:
              assign the coordinates of t_{i-1}^dr to t_i^dr
          determine z based on the coordinates of o from the tracking system and the office dimensions
        end
    end
  end
  T^total_start = the earliest arrival time to the office among all occupants
  T^total_end = the latest departure time from the office among all occupants
  TOD = T^total_end - T^total_start
  For t_i^dr in TOD
      if o is present:
        add 1 to n_o^(t_i^dr, d_i)
        calculate the occ_r based one n_o^(t_i^dr, d_i)
      else no change in the number of present occupants
   end
end
```

Figure 4-3 Pseudocode for data processing phases: zone and occupancy rate calculations

4.2.2.1 Occupancy Rate

The occupancy rate for time-step t, ($occ_r^{t,d}$), is the average occupancy rate for each day of a week based on the total number of weeks of the data collection (W). After collecting data for a certain period, the occupancy rate (%) of all zones within an office is calculated for each time-step (e.g., one minute) and for each day of a week (including weekends) according to Equation (4-1):

$$occ_r^{t,d} = \frac{\sum_{w=1}^{W}(\frac{n_o^{t,d}}{N_o})_w}{W} \times 100\% \tag{4-1}$$

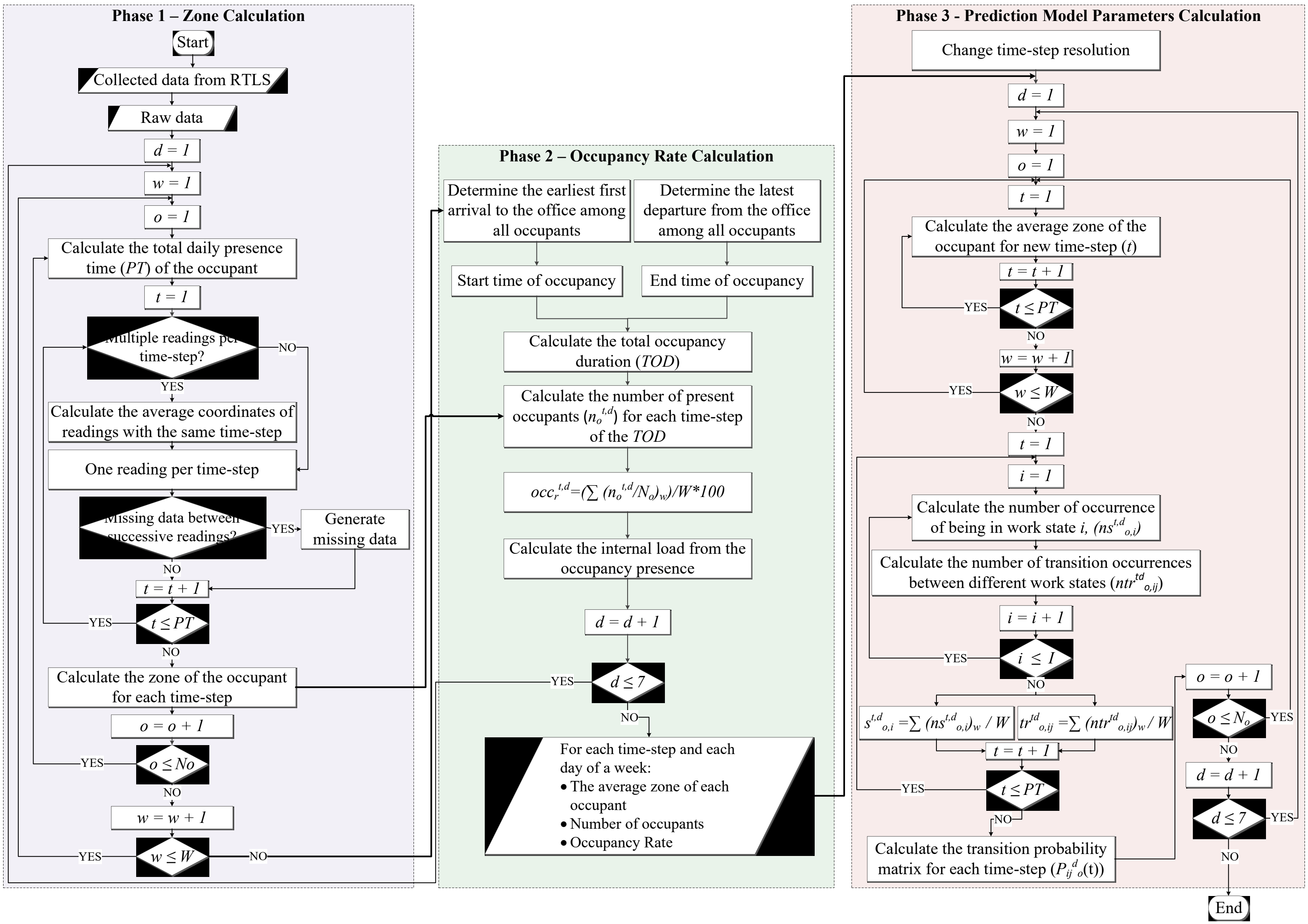


Figure 4-4 Occupant behavior analytics *(*data processing) phases

where $n_o^{t,d}$ is the number of present occupants at time-step t and day d, and N_o is the total number of occupants sharing the same open-plan office during day d.

4.3 Markov Chain Occupancy Prediction Model

In this research, the Markov chain technique is used for the analysis aiming to develop the probabilistic occupancy profiles. Since the occupants' movement among the zones inside and outside open-plan offices creates the occupancy profile, random mobility between different work states is assumed. This assumption allows for modeling the transitions among work states as a Markov chain process. Therefore, the next work state of the occupant only depends on his/her present state and some rules about the work states.

A Markov chain is a sequence of random variables with the Markovian property presented as (Serfozo, 2009):

$$\begin{aligned} P\{X_{t+1} = j | X_t = i, X_{t-1} = i_{t-1}, X_{t-2} = i_{t-2}, \dots, X_1 = i_1, X_0 = i_0\} \\ = P\{X_{t+1} = j | X_t = i\} = P_{ij}(t+1) \end{aligned} \tag{4-2}$$

where X_t is a random variable representing different occupants' work states, t is the time-step, and all states $i_0, i_1, \dots, i_{t-1}, i, j$ are nonnegative integers values $\in I = \{0, 1, 2, \dots\}$. $P_{ij}(t+1)$ shows the probability of transition from state i to state j at time $t+1$.

Knowing that the future state of the occupant depends on his/her current state, the transitions of states are defined in Markov matrices. Since the whole day is clustered into different time slots, as mentioned in Section 4.2, the probability of occurrence of different states varies with the time of the day; and consequently, the transition probability matrices are different for each of these time slots as illustrated in Figure 4-5. This figure shows the transition probabilities during the lunch break. For instance, if an occupant is going out of the office (at time *t*) for the lunch break at time *t+1*, there is a higher probability to either stay at lunch break or go back to his/her zone at time *t+2* and no probability to go to a short break. This makes the transition probabilities to be time-dependent. This type of Markov chain process is called an inhomogeneous Markov chain (Douc et al., 2004).

In the proposed inhomogeneous Markov chain model for prediction of space occupancy in multi-occupied offices, the states of the Markov chain are occupants' work states as shown in Table 4-1. This results in having 5×5 transition probability matrices independent of the maximum number of occupants in open-plan offices. Compared to methods that define transition probability matrices based on the number of occupants in a zone, (e.g. (Ai et al., 2014; Richardson et al., 2008; Han et al., 2012)), or methods that consider some restrictions regarding the movement of occupants between zones to reduce the order of transition matrices (e.g. (Chen et al., 2015)), using the proposed method significantly simplifies the calculation of transition probability matrices. Transition probability matrices are key parameters in Markov chain models and reducing their order has a high impact on the overall complexity of the Markov models, especially for inhomogeneous Markov chain models with a large number of transition matrices.

The output of the proposed inhomogeneous Markov chain model is the probabilistic profiles of each specific occupant. The work state of each occupant, his/her location and the total number of present occupants can be derived from these profiles at each time-step. Eventually, building energy-consuming systems are adjusted based on this information to reflect the variations in different occupants' daily profiles (Salimi et al., 2019).

4.3.1 Transition Probability Matrices

In the first step of determining the transition probability matrices, the PT of each occupant for different days of a week along with the distribution of being in different work states during the PT are deduced from the results of the Phases 1 and 2 of the data processing steps. Next, the transition probabilities between different work states are calculated. To do so, two parameters are required: (1) the percentage distribution of each work state for each occupant; and (2) the transition occurrences between different work states for each occupant. The transition probability matrix is then calculated using Equations (4-3) and (4-4):

$$P_{ij} = 1 - s_i + s_i \times tr_{ij} \qquad (if\ i = j) \tag{4-3}$$

$$P_{ij} = s_i \times tr_{ij} \qquad (if\ i \neq j) \tag{4-4}$$

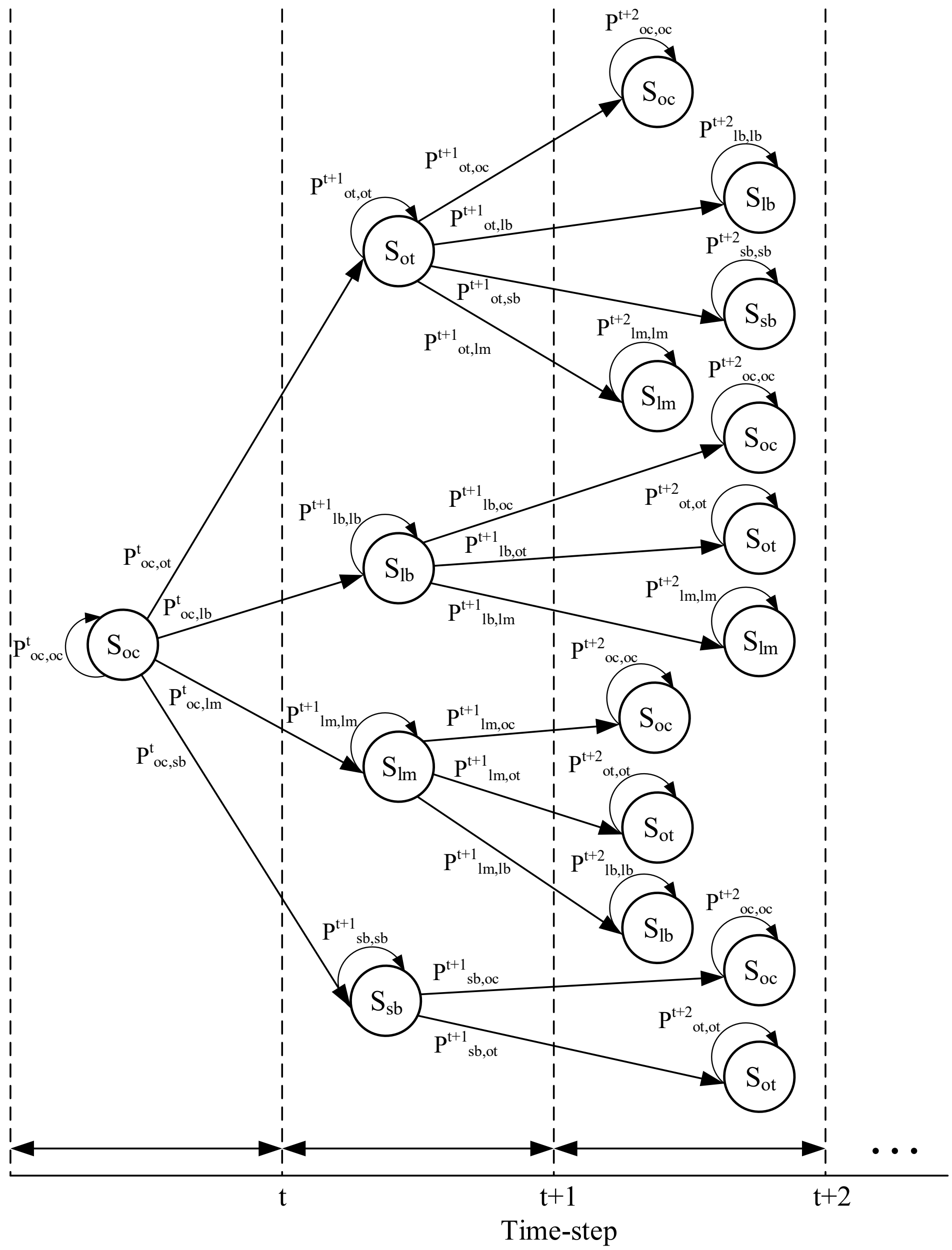


Figure 4-5 Sample Transition Process from Time *t* to *t+2* during Lunch Break

where s_i shows the probability of being in state *i,* which is calculated based on the percentage of occurrence of each work state during the monitoring period. The probability of transition occurrences from state *i* to state *j* is indicated by tr_{ij}. These formulas are inspired by the work of Yamaguchi et al. (2003). However, improvements are applied to their proposed formula. Firstly, the Markov chain is time-independent in their method. Secondly, they assumed constant numbers for the parameters s_i and tr_{ij}. In this study, the Markov chain and the parameters are time-dependent. In addition, the collected data regarding the actual occupancy of the open-plan office are used to define the parameters s_i and tr_{ij} with some enhancement in their calculation method as discussed below:

(1) For each occupant o ($o = 1,2, \dots, N_o$) and each day of a week d ($d = 1,2, \dots, 7$), the probability of being at work state i ($i = 1,2, \dots, I$, where I represents the maximum number of work states) at each time-step t, $s_{o,i}^{t,d}$, is obtained by counting the number of times of being in work state i ($ns_{o,i}^{t,d}$) divided by the total number of weeks over the monitoring period as shown below:

$$s_{o,i}^{t,d} = \frac{\sum_{w=1}^{W}(ns_{o,i}^{t,d})_w}{W} \tag{4-5}$$

This procedure results in personalized probability distribution graphs for each work state at different time-steps over each specific day of a week. The following condition should be considered for each time-step t and each occupant o when calculating the probabilities:

$$\sum_{i=1}^{I} s_{o,i}^{t,d} = 1 \tag{4-6}$$

(2) For each occupant o, the number of transition occurrences from state i to state j at each time-step t and for each day of a week d, $ntr_{o,ij}^{t,d}$, is obtained from the collected data. Then, the probabilities of transition occurrences are calculated over the monitoring period ($tr_{o,ij}^{t,d}$):

$$tr_{o,ij}^{t,d} = \frac{\sum_{w=1}^{W}(ntr_{o,ij}^{t,d})_w}{W} \tag{4-7}$$

In this study, the transition probability matrix for each time-step and for each day of a week is then calculated for each occupant (${P_{ij}}_o^d(t)$) using Equations 8 and 9:

$${P_{ij}}_o^d(\mathrm{t}) = 1 - s_{o,i}^{t,d} + s_{o,i}^{t,d} \times tr_{o,ij}^{t,d} \qquad (if\ i = j) \tag{4-8}$$

$${P_{ij}}_o^d(\mathrm{t}) = s_{o,i}^{t,d} \times tr_{o,ij}^{t,d} \qquad (if\ i \neq j) \tag{4-9}$$

Considering five states of the transition probability matrix, this matrix has a dimension of $5 \times 5 \times 288 \times 7$ using a 5-minute time-step for one day (i.e., 288) and one matrix for each day of a week (i.e., 7 days). The procedure for finding the transition probability matrix is demonstrated in Figure 4-6.

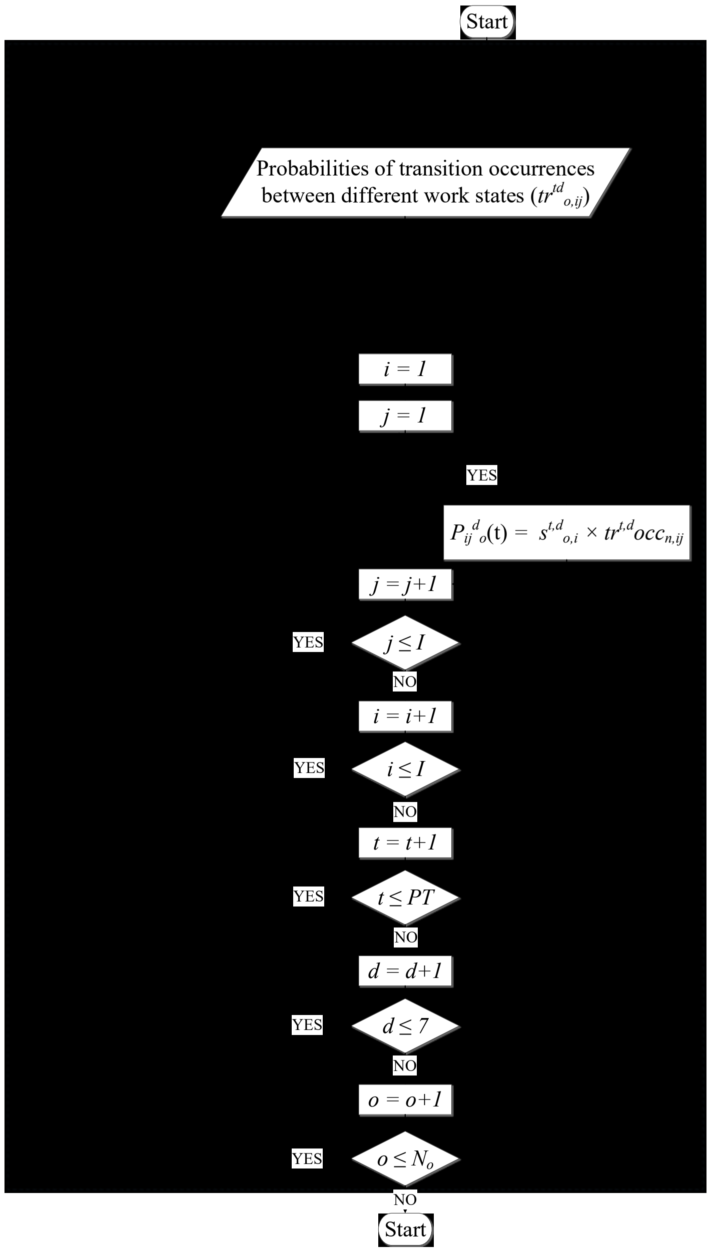


Figure 4-6 Markov chain transition matrix flowchart

In the next step, Probability Density Function (PDF) for each time-step t can be deduced from each row of the Markov transition matrix. Further, the Cumulative Distribution Function (CDF) is derived from the PDF for each time-step. The CDF is a histogram of five bins corresponding to the five work states. Each bin shows the probability at which a value of that bin can be randomly selected. Using the Inverse Function Method (IFM) gives the estimation of the work state for the next period $(t+1)$. The IFM works by inverting the CDF of the parameter of interest. It randomly generates a number between 0 and 1 using a uniform distribution. The random number determines which bin is going to be selected for the parameter of interest using the CDF. Figure 4-7 illustrates these steps.

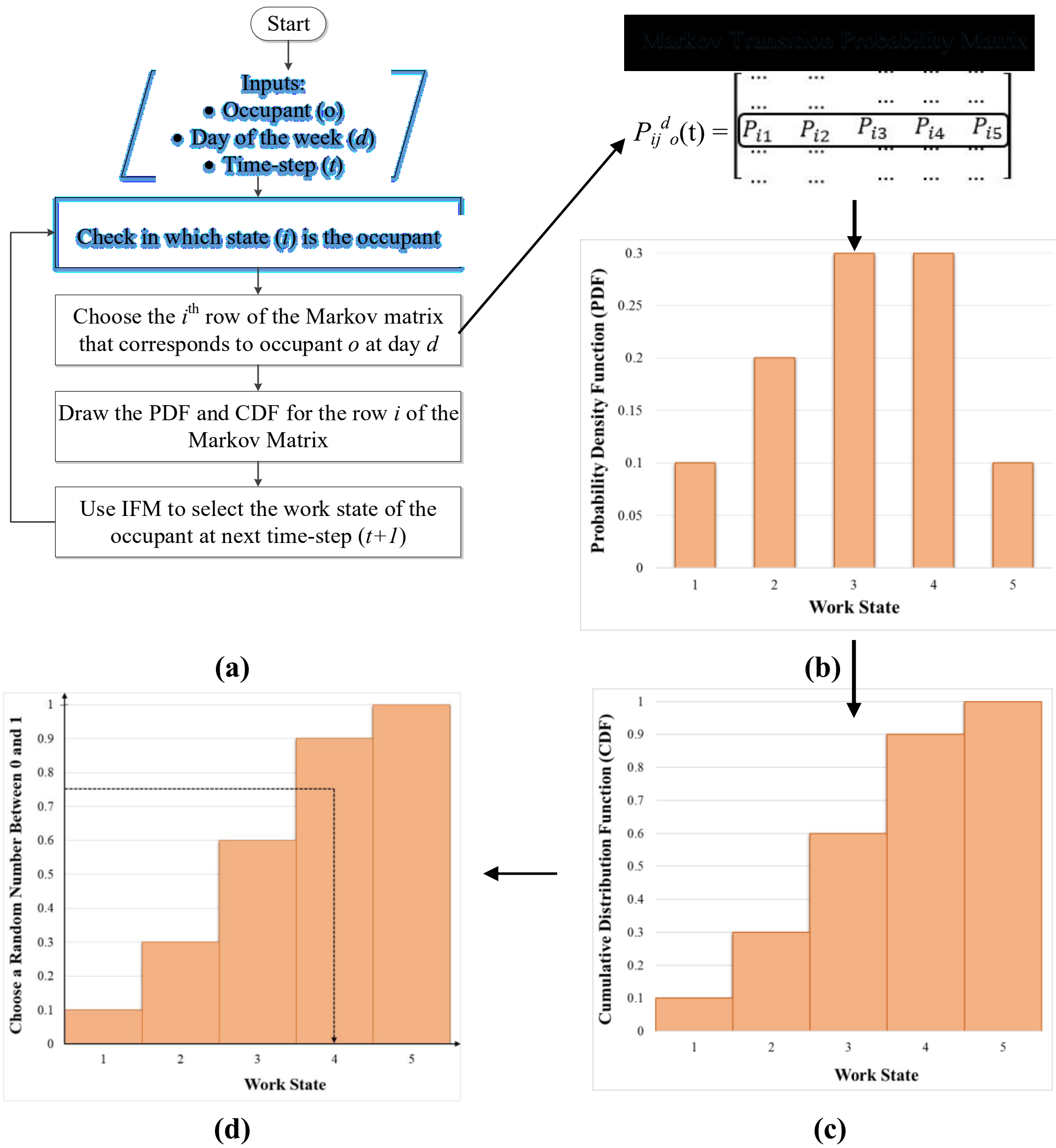


Figure 4-7 Work state estimation flowchart; (b) PDF generation; (c) CDF generation; (d) Generation of series of work states using IFM

4.4 Implementation and Case Study

Figure 4-8 shows the picture of the case study location (a research laboratory) and the monitoring system set up along with the office layout. There are six occupants assigned to the research laboratory. In this case study, the occupancy data were collected every second using Bluetooth Low Energy (BLE) (also known as Bluetooth 4.0 or Bluetooth Smart) for one month. The BLE-based monitoring system used in this research (i.e., Quuppa Intelligent Locating System™) is able to track the latest smartphones and BLE devices with an accuracy of 20-50 *cm* (Liu, 2017; Analytics, 2015). Based on the measurements made, the size of the monitored office is 5.0 m×7.0 m×3 m. In order to get the required data for the prediction model, it is important to know whether the occupant is at zone 1, 2 or 0 (which is the outside of the office) as shown in Figure 4-8. According to the dimensions of each zone (i.e., 5.0 m×3.5 m), the accuracy of 20-50 *cm* is precise enough for the purpose of this study.

Quuppa system uses the AoA approach to calculate the position of different objects (e.g., people, equipment, etc.) as discussed in Section 4.2.1. This system offers many advantages including long tag battery lifetime, compatibility with standard mobile devices, and the ability to carry sensor data alongside the positioning data (Quuppa Intelligent Locating System™, 2016).

As shown in Figure 4-8-(b), four locators are used in this study to accurately monitor the occupants and their movement. According to (Quuppa, 2017), distances of 6-10 meters between indoor locators are convenient for a good coverage. In our case, locators are placed with the distances between them less than 7 meters. In addition, the coverage quality estimate is checked and demonstrated in Figure 4-9. As shown in this figure, the red color represents bad quality and green color represents good quality. Thus, the coverage quality of four locators in the room is good for tracking. In the case of having larger offices, more sensors are required to accurately cover the whole space in order to collect precise occupancy data. Having enough number of locators with distances within the suggested range, the same accuracy of 20-50 cm is achievable.

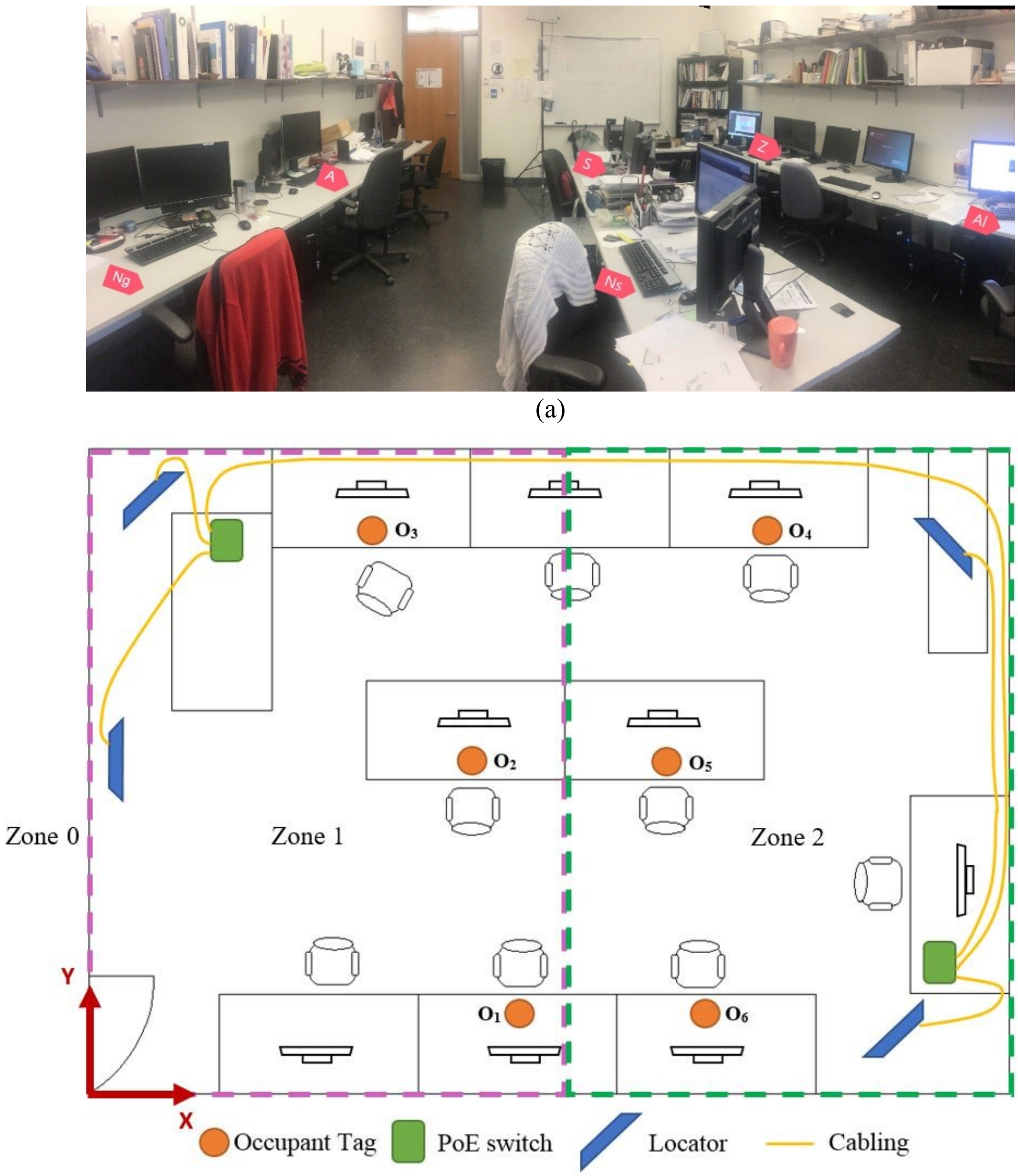


Figure 4-8 (a) Case study location (graduate research lab); (b) Monitoring system setup and office layout

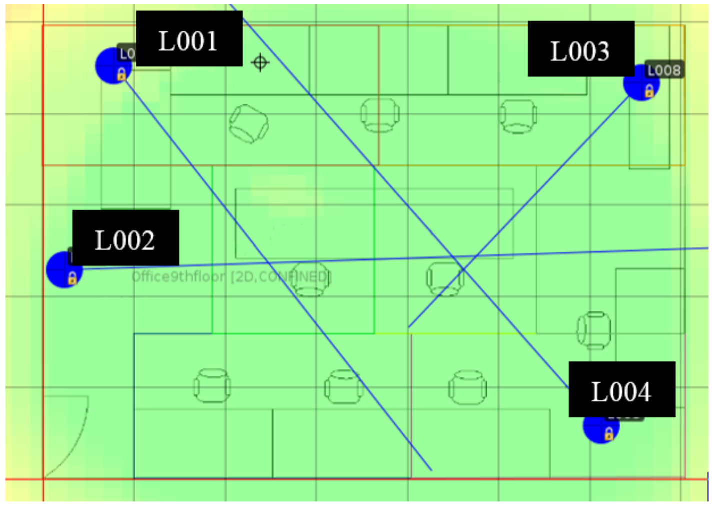
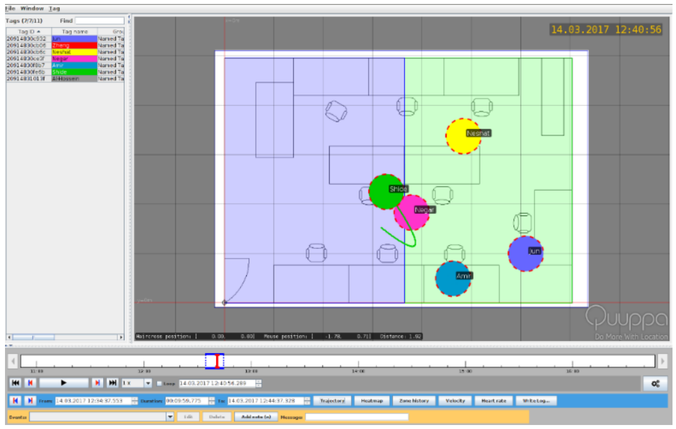


Figure 4-9 Quality Estimate of the Locators Coverage (L001-L004 are the four locators used in the monitoring process)

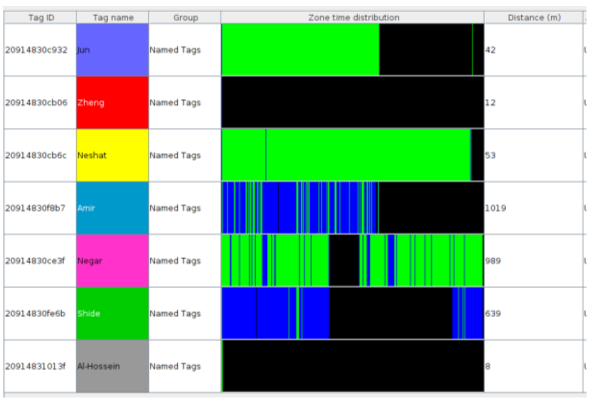
4.4.1 Visualization of Recorded Data in Quuppa

After collecting the data, the QDP (Quuppa Data Player) allows reviewing the recorded data with respect to the paths occupants followed, the percentage of the time that each occupant spends in each zone, and the heat map during the recorded time. The occupants' paths in Figure 4-10(a) show all the movements of the occupants, which is the replay of the passed monitoring time. As can be seen in Figure 4-10(b), the occupancy-time chart indicates the length of the time that the occupants spend in each zone. The heat map in Figure 4-10(c) can be used to define the most occupied area of the space.

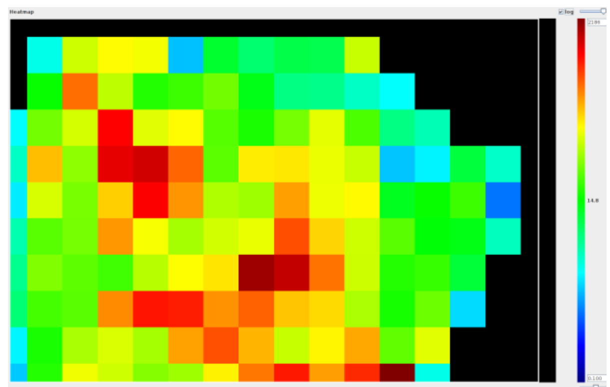
In addition, the scattered plots are created to visualize the distribution of the occupants' positions. For example, occupants 1 and 2 are assigned to Zone 1 and Zone 2, respectively. As Figure 4-11 shows, most of the time, the occupants are in their zones, but they also interact with other occupants or appliances in other zones. The scatter plots can only provide the distribution of the occupants' movements, which are related to the *Who* and *Where* questions, but the occupancy schedule cannot be shown.



(a) Occupants' paths

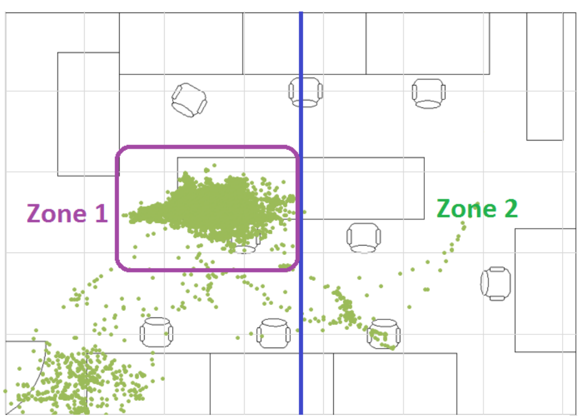


(b) Occupancy-time proportion for each zone

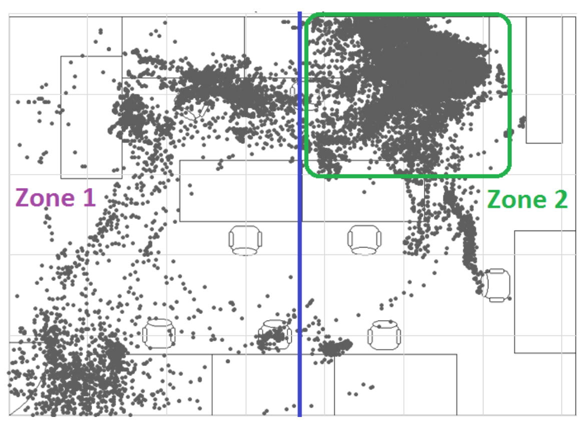


(c) Heat map

Figure 4-10 Data visualization enabled by the Quuppa system



(a) Occupant 1



(b) Occupant 2

Figure 4-11 Scattered plots of the movement distribution of the occupants

4.4.2 Occupancy Probabilistic Profiles

The probabilistic profile of each occupant shows the probability of the occupant's presence at a certain time of the day at a specific location. This profile can be used for predicting the status of the occupants and adjusting building operational systems in advance to save energy as well as to satisfy the occupants' indoor environment comfort levels.

In this study, the test was run for one month and since the collected data from the monitoring system could be used for different purposes with different levels of accuracy, the BLE system monitored occupants with high resolution (i.e., each second). Collecting the occupancy data with the high resolution of one second generated about 250 MB of the raw data in total. Figure 4-12 shows the distribution of the size of the collected data over the one-month period of the data collection. It takes some time to polish this raw data, such as producing the missing data or removing duplications in the collected raw data, as explained in Section 4.2.2.

All the data processing phases and the development of the occupancy prediction model are performed on a desktop computer with properties as Intel Xeon CPU X5550 @ 2.67 GHz, 6 GB Random Access Memory (RAM), and running Windows 7 Professional Dell computer.

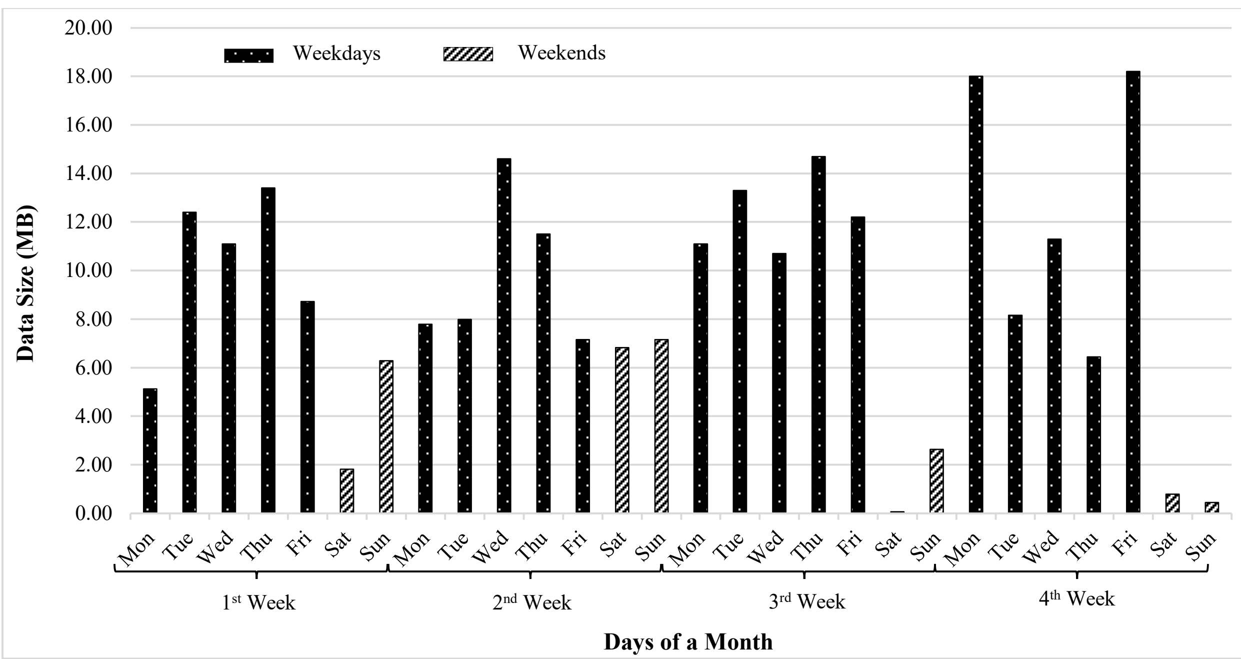


Figure 4-12 Size of the collected data per day

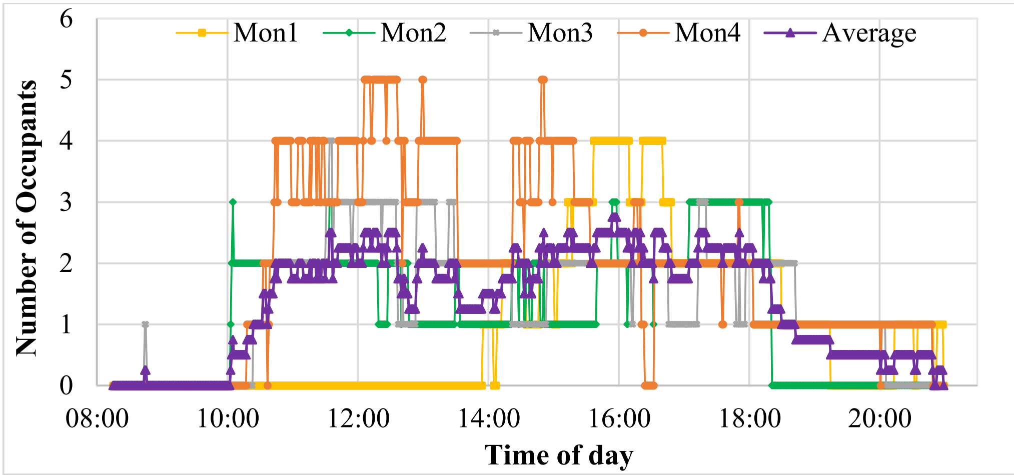
The data processing phases are used to produce the input data to the prediction model with the desired time-step. The required computational time is about four hours for one month of the collected data. However, the high granularity of one second is not required for building energy management. Therefore, the occupants' zones are calculated every five minutes according to

Section 4.2.2. During a five-minute time-step, the final selected zone for that time-step will be the zone in which the occupant spent more minutes. In this study, the number of defined zones is equal to $\boldsymbol{M}$ plus one zone for the outside of the office. For instance, three zones are considered for a shared office with zones 1 and 2, being within the office, and one zone for outside of the office (i.e., zone 0).

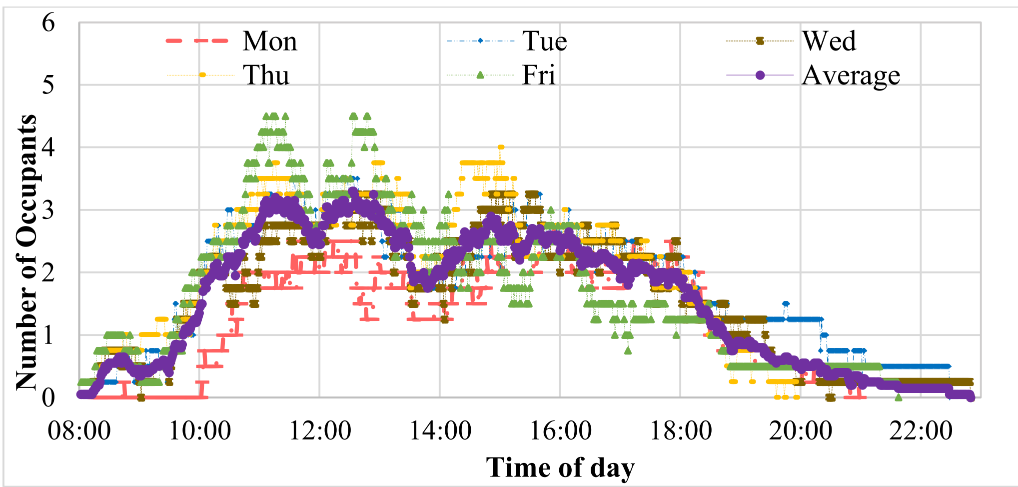
After determining the occupants' zones for each time-step of PT, the average number of occupants present at each zone of the room is calculated for each day of a week for a five-minute time-step resolution. The average occupancy rate for each day of a week (average of four series of data) is also calculated using Equation (4-1). Figure 4-13 illustrates the results for Mondays. The same results are obtained for other days of the week, which are not included here due to space limitation. This figure shows the number of present occupants at the office level for four Mondays in a month. The average occupancy number for one month is also shown in this figure. Furthermore, the average occupancy numbers for weekdays and weekends are illustrated in Figure 4-13. As it was expected, the occupancy rate of the office is much lower during weekends. However, it is important to know that the office is always occupied for several hours during the weekends.

4.4.3 Validation of the Occupant Behavior Analytics Method

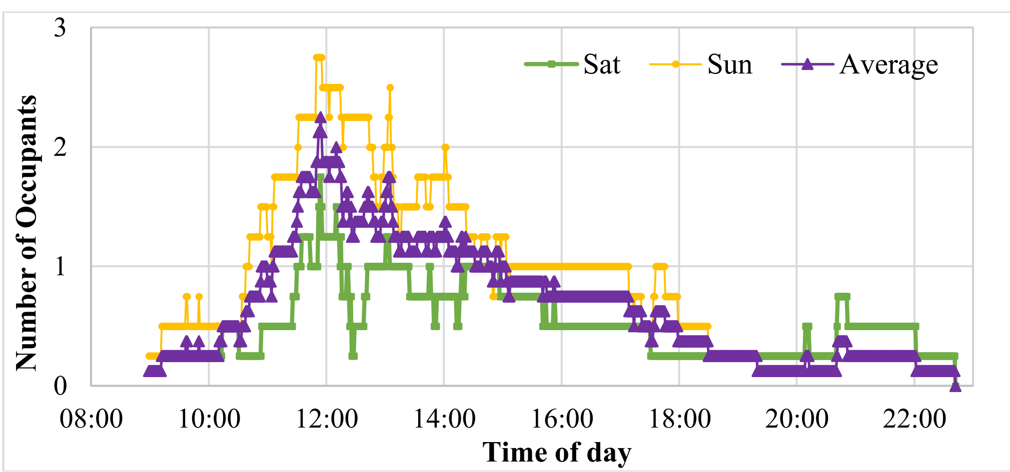
The proposed data processing method is validated by comparing the obtained profiles using the proposed method with those obtained from the ground truth data. A check table is created to collect the ground truth data from April 07 to May 17, 2017, and from May 23 to June 10, 2017. The table includes the time of occupants' first arrival, their lunch break, and the last departure. After logging the information, the PDF of the events of importance along with their CDF, such as the first arrival to the office, are created to determine the actual range of their occurrences (Figure 4-14 and Figure 4-15). As shown in these figures, the majority of the first arrival event has occurred between 08:15 to 11:30 am. Table 4-2 shows the start and end times of these events based on the ground truth data as well as ranges that are used for data processing at the office level. As discussed in Section 4.2, five typical work states are considered in office buildings (Table 4-1).



(a)



(b)



(c)

Figure 4-13 Variation in occupancy number for (a) Mondays; (b) Weekdays; and (c) Weekends

In order to process the raw data to reflect these work states, the ranges derived from ground truth data, as a basis for data processing, are broken down to smaller ranges. In addition, a window of $\pm$15 minutes is considered at the start and end times of these ranges to give more diversity to these time slots. This means that 15 minutes is deducted from the first arrival time and 15 minutes is added to the end time of each range (fifth and sixth columns in Table 4-2). Furthermore, the same time ranges showing different work states' durations are created for each occupant to have more accurate time ranges that are specific to each occupant's pattern of presence for each day of a week as demonstrated in Table 4-2.

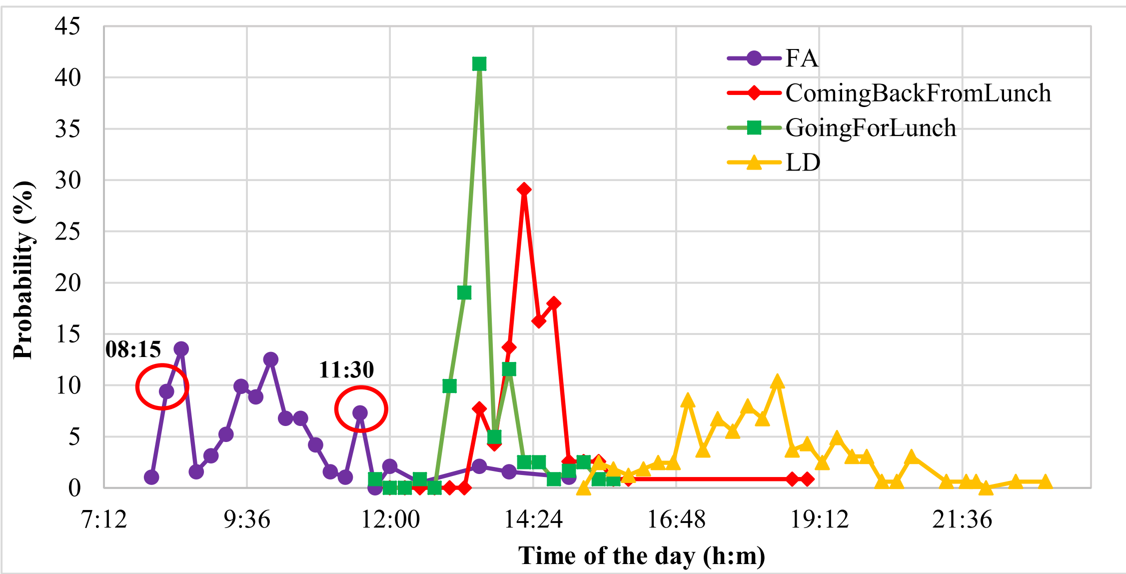


Figure 4-14 Probability Distributions of Events of Importance during a Day

4.4.3.1 Comparison with the Ground Truth Data

The comparison between the occupant behavior analytics method and the ground truth data is displayed in Table 4-3. There are some cases that differences between the arrival time and the results of the occupant behavior analytics method are found. The review of the raw data in these cases shows that the occupants arrived and left the room after a short stay (less than five minutes). This results in a delay for the processing methods to capture the first arrival of the occupants, such as the first arrival of occupant *Ng* on May 1st that is not captured correctly by the occupant behavior analytics method. On the other hand, the occupant behavior analytics method could catch the departure times. Hence, it can be concluded that the daily profiles are in accordance with the ground truth data.

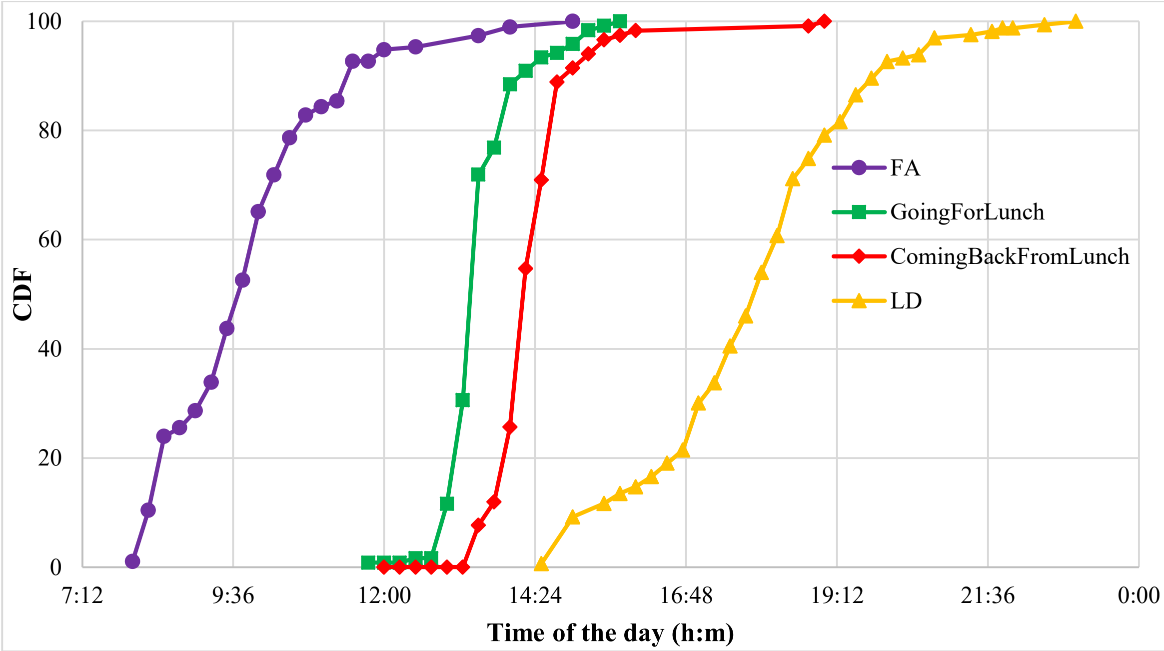


Figure 4-15 Probability distributions of events of importance during a day

Table 4-2 Examples of start and end times of events of importance at the office and at occupant level

Level	**Events**	**Ground Truth**		**Data Processing**	
		Start Time	**End Time**	**Start Time**	**End Time**
Office	First arrival	08:15	11:30	08:00	11:45
	Before lunch	11:30	13:30	11:45	13:45
	Lunch break	13:30	16:00	13:45	16:15
	After lunch	16:00	18:15	16:15	18:30
	Last departure	18:15	20:45	18:30	21:00
Occupant	First arrival	09:15	10:15	09:00	10:30
	Before lunch	10:15	13:30	10:30	13:45
	Lunch break	13:30	14:15	13:45	14:30
	After lunch	14:15	17:30	14:30	17:45
	Last departure	17:30	20:45	17:45	21:00

Table 4-3 Validation of the occupant behavior analytics method

		Arrival			Departure		
Date	**Occupant**	**Ground Truth**	**Data Processing Method**	**Error (%)**	**Ground Truth**	**Data Processing Method**	**Error (%)**
5.01	Ng	08:30	09:48	13	18:15	18:19	0
	A	09:30	10:04	6	18:15	18:21	1
	S	11:30	10:05	-14	18:15	18:21	1
	Ns	*NP	15:53	–	18:15	18:19	0
	Al	*NP	*Ab	0	*NP	*Ab	0
	Z	*NP	*Ab	0	*NP	12:39	–
5.02	Ng	08:30	08:38	2	19:00	18:48	-1
	A	09:30	10:05	6	*NP	13:47	–
	S	09:00	09:07	1	15:50	15:50	0
	Ns	09:30	10:04	6	17:00	16:47	-1
	Al	*NP	*Ab	0	*NP	*Ab	0
	Z	*NP	*Ab	0	*NP	*Ab	0
5.03	Ng	08:15	08:14	0	19:30	19:26	0
	A	09:30	09:38	1	23:00	22:50	-1
	S	10:45	13:57	23	19:30	19:32	0
	Ns	10:30	10:43	2	*NP	17:00	–
	Al	*NP	*Ab	0	*NP	*Ab	0
	Z	*NP	13:17	–	*NP	13:29	–
5.04	Ng	08:30	08:22	-2	*NP	19:28	–
	A	08:45	08:22	-5	*NP	15:11	–
	S	09:45	10:14	5	17:45	17:57	1
	Ns	10:30	10:51	3	18:30	18:39	1
	Al	*NP	*Ab	0	*NP	*Ab	0
	Z	*NP	*Ab	0	*NP	*Ab	0
5.05	Ng	08:00	08:02	0	19:00	18:50	-1
	A	10:30	10:43	2	*NP	15:52	–
	S	10:45	10:47	0	17:45	17:46	0
	Ns	10:45	10:45	0	*NP	14:37	–
	Al	*NP	*Ab	0	*NP	*Ab	0
	Z	*NP	11:10	–	*NP	11:45	–
5.06	Ng	11:15	11:20	1	16:00	15:45	-2
	A	10:45	10:53	1	*NP	12:23	–
	S	*NP	*Ab	0	*NP	*Ab	0
	Ns	*NP	*Ab	0	*NP	*Ab	0
	Al	11:30	11:35	1	17:30	17:30	0
	Z	*NP	Ab	0	*NP	*Ab	0
5.07	Ng	*NP	09:12	–	*NP	17:46	–
	A	*NP	*Ab	0	*NP	*Ab	0
	S	*NP	*Ab	0	*NP	*Ab	0
	Ns	*NP	*Ab	0	*NP	*Ab	0
	Al	*NP	09:50	–	*NP	14:42	–
	Z	*NP	*Ab	0	*NP	*Ab	0

*NP: Not Provided
*Ab: Absent

4.4.4 Occupancy Prediction Results using the Markov Chain Model

To estimate the occupancy profiles using the probabilistic inhomogeneous Markov chain occupancy prediction model, all states are labeled to show the transition probabilities from one state to another according to Table 4-1. For example, the transition probability from work state 1 to 3, which is leaving the office for lunch break, is presented by ${P_{oc,lb}}_{o}^{d}(\mathrm{t})$. Thus, the transition probability matrix of ${P_{ij}}_{o}^{d}(\mathrm{t})$ can be shown as follows:

$$
{P_{ij}}_{o}^{d}(\mathrm{t}) = \begin{bmatrix} P_{oc,oc} & P_{oc,ot} & P_{oc,lb} & P_{oc,sb} & P_{oc,lm} \\ P_{ot,oc} & P_{ot,ot} & P_{ot,lb} & P_{ot,sb} & P_{ot,lm} \\ P_{lb,oc} & P_{lb,ot} & P_{lb,lb} & P_{lb,sb} & P_{lb,lm} \\ P_{sb,oc} & P_{sb,ot} & P_{sb,lb} & P_{sb,sb} & P_{sb,lm} \\ P_{lm,oc} & P_{lm,ot} & P_{lm,lb} & P_{lm,sb} & P_{lm,lm} \end{bmatrix} \tag{4-10}
$$

As discussed in Sections 4.1 and 4.2.2, different resolution levels are required for controlling different building systems. For instance, a higher level of resolution is needed to apply lighting control strategies, which improve the comfort level. However, considering the required lag time for HVAC systems to adjust the indoor temperature to a specified target set-point, a lower level of resolution may not lead to a significant thermal discomfort. As a result, two different prediction time-steps are defined to determine occupancy predictions for lighting and HVAC systems control.

Five-minute prediction time-step is considered to predict the office occupancy pattern and accordingly adjust the lighting system. While this time-step changes to 30-minute prediction time-steps to control the HVAC system. Having the occupants' zones for every five-minute time interval, the distribution of the time being spent in the office's zones and outside is determined for each day of a week (i.e., $ns_{o,i}^{t,d}$). After calculating the number of transition occurrences (i.e., $ntr_{o,ij}^{t,d}$), the transition matrices corresponding to each time-step of each day of a week would be calculated using the average values of $tr_{o,ij}^{t,d}$ and $s_{o,i}^{t,d}$ for that specific day of the week throughout the whole month. For instance, the transition matrix of occupant o_5 on Mondays at 02:40 pm is shown below:

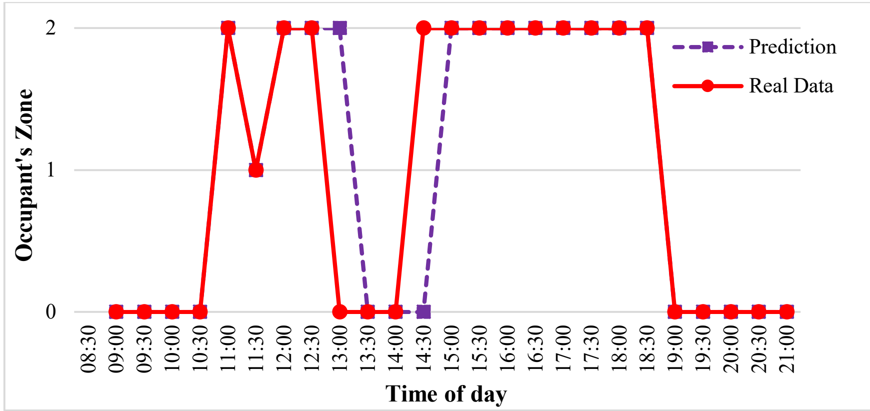
$$P_{ij_{o_5}}^{Mon}(02{:}40\text{ pm}) = \begin{bmatrix} 0.8125 & 0 & 0.0625 & 0 & 0.125 \\ 0 & 0 & 0 & 0 & 0 \\ 0.0625 & 0 & 0.8125 & 0 & 0.125 \\ 0 & 0 & 0 & 0 & 0 \\ 0.125 & 0 & 0.125 & 0 & 0.75 \end{bmatrix} \quad (4\text{-}11)$$

In this matrix, a row of zero probabilities happens when the $s_{o,i}^{t,d}$ is zero. In these cases, since the probability of being in state, i is zero at that specific time-step, it is not possible to have probabilities of state transitions.

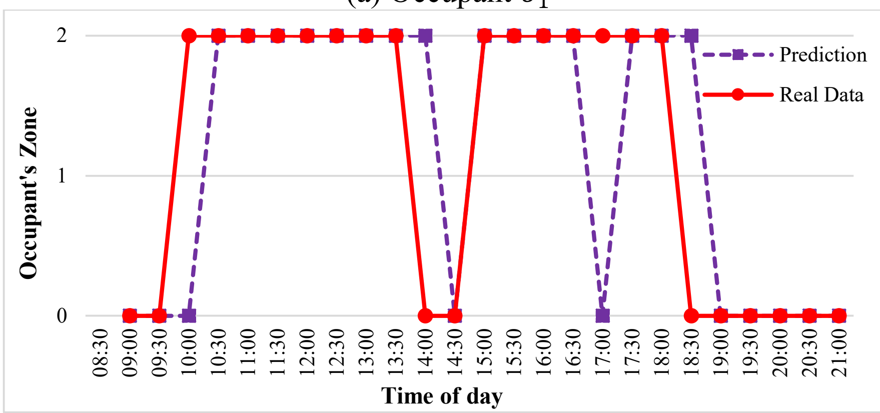
4.4.5 Validation of the Probabilistic Inhomogeneous Markov Chain Occupancy Prediction Model

To validate the performance of the probabilistic inhomogeneous Markov chain occupancy prediction model, the actual occupancy for different days of a week are compared to those of resulted from the prediction model. The comparison between the occupancy profiles resulting from the prediction model, to be used for the purpose of lighting control (i.e., five-minute time-step prediction), and the real data is illustrated in Figure 4-16 for occupants o_1 and o_6. This figure shows the zone of each occupant for each five-minute time-step. Comparing the prediction results and the actual occupancy patterns show the high accuracy of the prediction model (92% and 84% for occupants o_1 and o_6, respectively) in capturing the variations in occupants' zones.

As mentioned in the previous section, two different time-steps are defined to determine occupancy predictions for lighting and HVAC systems' control. In the case of using occupancy prediction to control the HVAC system, the initial state of occupancy is determined using the real-time collected data. Then, the probabilistic inhomogeneous Markov chain prediction model predicts the occupancy pattern for the next 30 minutes using the transition probability matrices and the IFM method as shown in Figure 4-7.



(a) Occupant o_1



(b) Occupant o_6

Figure 4-16 Comparison of predicted and the real occupancy profiles-lighting control purposes (Mondays)

Since the ${P_{ij}}_o^d(\mathrm{t})$ matrices are derived based on five-minute time-steps, the prediction model should be run six times to produce the occupancy pattern for the next 30 minutes. Then, the actual occupancy state is read again from the occupancy sensors to restart the prediction procedure. This update improves the accuracy of the prediction model by avoiding the accumulation of errors happening at each five-minute time-step. The prediction process is then repeated for the next 30 minutes and this loop is continued till the end time of occupancy *PT*. Figure 4-17 demonstrates this procedure. The same method is applicable in the case of using occupancy prediction to control the lighting system with the difference of changing the 30-minute to five-minute time-steps. Thus, the prediction model is only run once in this case since the time-step of having the ${P_{ij}}_o^d(\mathrm{t})$ matrices and prediction time-step are identical.

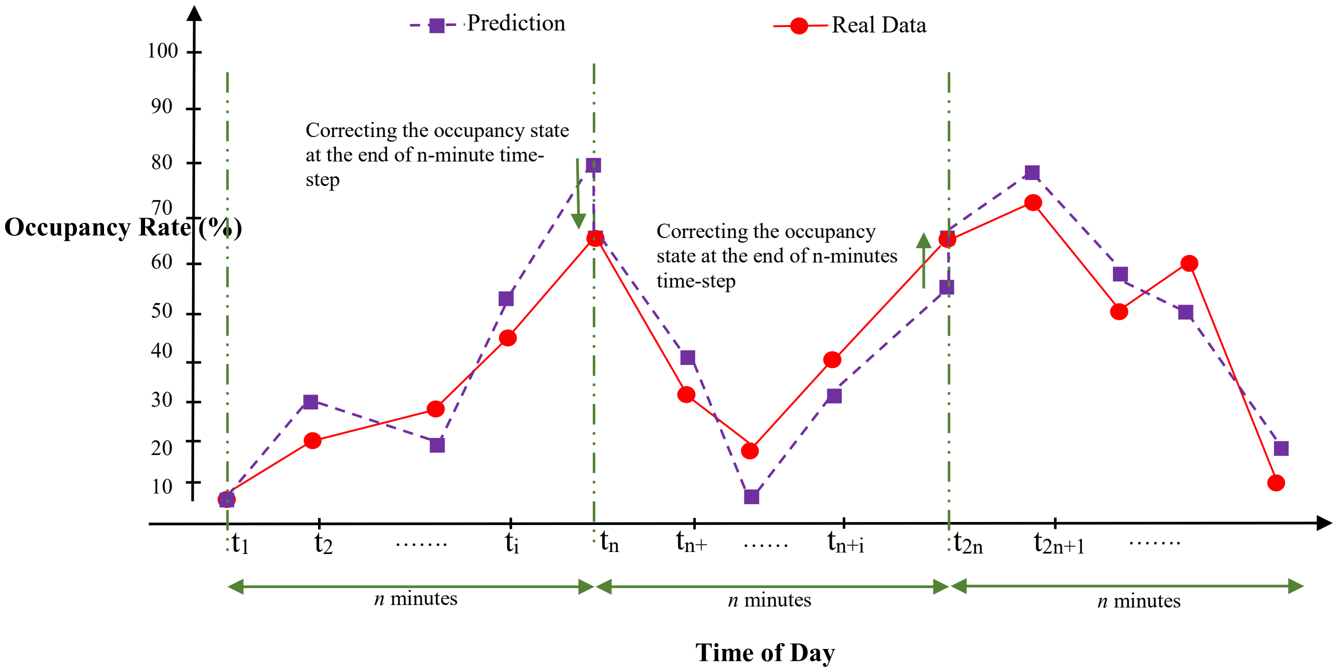
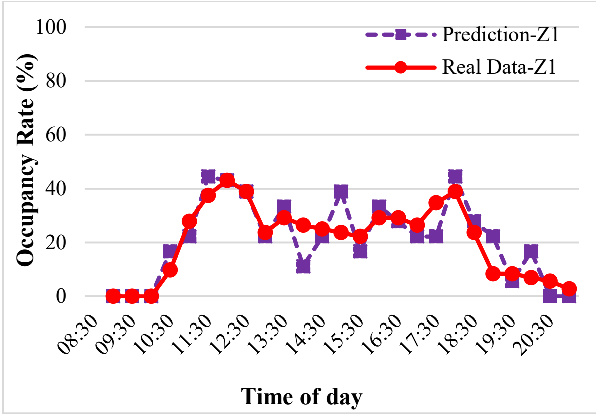


Figure 4-17 Occupancy prediction process and updates

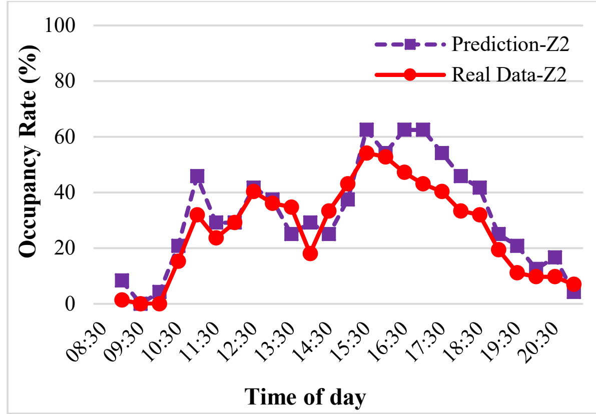
The proposed prediction model is an adaptive model that evolves and improves itself over time. The adaptive probabilistic occupancy model determines the HVAC and lighting systems' settings at the beginning of each day based on the collected data. In this manner, the HVAC system starts working half an hour before the start time of occupancy to reach the desired temperature by the time the first occupant arrives at the office. The lighting system will be turned on at the time of arrival of the first occupant. Using the prediction model, the level of occupancy is estimated for the next time-step and the HVAC and lighting systems are accordingly adjusted. This procedure continues for each time-step during the day (as specified for the control purpose) until the estimated end time of occupancy is reached. However, there may be some cases that the real occupancy does not follow the predicted one. This could happen when an unexpected occupancy happens when the prediction model estimates a vacancy for space. For instance, one occupant can arrive earlier than the time she/he was expected to start working. In these cases, there would be a switch to the real-time operation of building systems where the sensors detect the unexpected occupant and turn on the light automatically. Thus, real-time occupancy detection and control give an indication of the wrong estimation of occupancy. In such cases, an update is sent to the adaptive prediction model to adjust itself by correcting the current state of the occupancy. Using newly collected occupancy data helps the model to capture changes in occupancy space utilization patterns, especially in the case of open-plan offices with varying occupancy (e.g., common labs). In this type of offices, occupancy data should be collected over a shorter period, and more frequent updates are required to reflect changes in occupancy patterns. This would help the model to evolve and become more precise in predicting office occupancy. Therefore, the idea of frequently updating the occupancy prediction model can improve the reliability of the model for future predictions. However, the data collection period can be longer for offices with fixed occupancy, such as research labs; since there would not be many variations in the space utilization patterns. Using different data collection periods and frequent updates make the proposed prediction model more general for different types of open-plan offices.

Having the occupancy profile prediction for each occupant results in developing the occupancy rate prediction for each zone. After calculating the ${P_{ij}}_o^d$(t) matrices, the number of present occupants in each zone can be predicted. Using Equation 1, the occupancy rate at the zone level is calculated and the results are used to control the lighting system. Since the prediction model demonstrated the same performance for different days of a week, Figure 4-18 shows the occupancy rates only for Mondays.

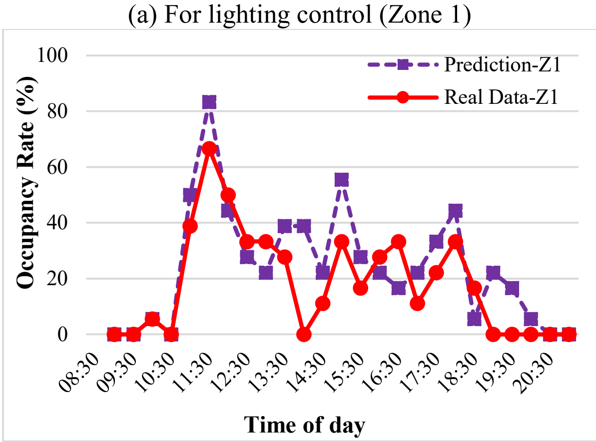
The same comparison was made between the occupancy rates resulted from the prediction model to be used for the purpose of HVAC system control (i.e., 30-minute time-step prediction), and the real data as also illustrated in Figure 4-18. As shown in this figure, the prediction model captures the real behavior of occupants at the zone level. The prediction model is able to accurately estimate the location of occupants at most periods of data collection during the day, which shows that the overall performance of the prediction model is satisfactory. Table 4-4 shows the performance measurement of the prediction model using the coefficient of determination (known as R^2) for different cases. The values of R^2 when using the proposed prediction model for the lighting control are 0.8 and 0.92 for zones 1 and 2, respectively. This would result in having 0.86 on average for this parameter (86%) for the application of lighting system control. The same method is used to calculate the average value of R^2 when using the proposed prediction model for the control of the HVAC system (68%). These values indicate the high accuracy of the prediction model in imitating the real occupancy patterns of the open-plan office.



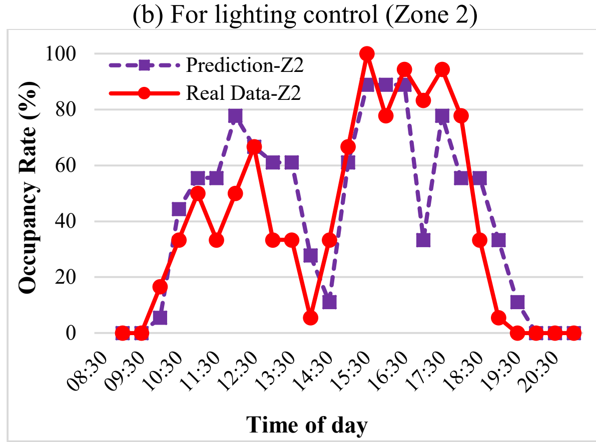
(a) For lighting control (Zone 1)



(b) For lighting control (Zone 2)



(c) For HVAC control (Zone 1)



(d) For HVAC control (Zone 2)

Figure 4-18 Comparison of predicted and the actual occupancy rates (Mondays)

Table 4-4 Performance measurement of the prediction model

Use of occupancy prediction	**Level of prediction**	R^2
Lighting control	Zone 1	0.8
	Zone 2	0.92
HVAC Control	Zone 1	0.65
	Zone 2	0.7

4.5 Summary and Conclusions

In this chapter, the occupancy modeling (i.e., occupants' profiles) has been further enhanced using a probabilistic inhomogeneous Markov chain prediction model based on real occupancy patterns data. The main contributions of this research are: (1) developing a method for extracting detailed occupancy information with varying time-steps from collected RTLS occupancy data. This method can capture different resolution levels required for the application of intelligent, occupancy-centered local control strategies of different building systems; (2) developing a new adaptive probabilistic occupancy prediction model based on the extracted occupancy information; and (3) developing time-dependent inhomogeneous Markov chain occupancy model, which distinguishes the temporal behavior of different occupants within an open-plan office.

The proposed prediction model is an adaptive model that evolves and improves itself over time. By frequently updating the occupancy prediction model whenever an unexpected occupancy happens, the model captures changes in the occupancy space utilization patterns and becomes more precise in predicting the office occupancy.

Having the occupancy profile prediction for each occupant results in developing the occupancy rate prediction at the zone level. The comparison between the occupancy profiles resulting from the prediction model and the actual profiles showed that the prediction model was able to capture the actual behavior of occupants at occupant and zone levels. The prediction model can accurately estimate the location of occupants at most periods of data collection during the day. High accuracy (86% and 68% on average for the purpose of the lighting and HVAC systems control, respectively) of occupancy patterns prediction also indicates the acceptable performance of the prediction model in capturing the temporal behavior of different occupants working in the same open-plan office.

Although the overall performance of the prediction model was satisfactory, it may not capture variations in occupancy patterns that may happen after the data collection period, especially in the case of open-plan offices with varying occupancy. This limitation could be solved by collecting occupancy data for a longer period of time and frequently updating the prediction model whenever a real-time occupancy detection and control happened to consider changes in the space utilization patterns.

There is a privacy issue when the occupants' identities are used to have detailed occupancy information. However, this issue can be resolved by anonymizing the occupants' data through defining occupancy profiles per zone. In addition, having this type of data could be vital for other purposes, such as emergency and safety. Informing the monitored occupants about all the benefits coming from using the real-time monitoring system for a reasonable period could also be helpful to solve this issue.

CHAPTER 5 SENSITIVITY ANALYSES OF THE OCCUPANCY PREDICTION MODEL

5.1 Introduction

An Important indicator for evaluating a building energy performance is its occupancy information including occupants' presence data and their behavior. Considering that more than 80% of buildings' energy consumption occurs during their operation phase makes occupancy a crucial element in assessing the energy usage of buildings (Liang et al., 2016). One of the accepted methods for estimating buildings' energy consumption is building simulation (Ioannou and Itard, 2015). There are different simulation tools in the market for analyzing energy performance of buildings, such as DOE-2 (DOE, 2016), EnergyPlus (EnergyPlus, 2015), IES-Virtual Environment (IES, 2019), ESP-r (ESRU, 2012), and TRNSYS (TRNSYS, 2013). Despite some minor differences between these tools, building parameters and occupancy information are common inputs among them. Building simulation tools are mature in terms of incorporating proper building parameters in energy analysis. However, they have some shortcomings with regard to occupancy data, which could cause inaccurate prediction of building energy performance (Clevenger and Haymaker, 2006). This makes the occupants-related parameters a driving factor causing large discrepancies in the building energy usage even between similar buildings with the same characteristics. According to (Ioannou and Itard, 2015), these differences range from 30% up to 100% in some cases (Soebarto and Williamson, 2001; Guerra-Santin and Itard, 2012; Majcen et al., 2013a-b).

The ability of simulation tools to accurately estimate buildings' energy usage close enough to their actual use, hence, is dependent on the accuracy of the provided input data. This makes the performance of the energy simulation models sensitive to the input parameters. In terms of sensitivity of the simulation models to the physical parameters of buildings, many studies have been conducted focusing on these parameters (Li et al., 2016; Zhang et al., 2017; Delgarm et al., 2018; Gagnon et al., 2018; Yip et al., 2019; Tian et al., 2019). Furthermore, the impact of changes in occupancy-related parameters on the performance of energy models has been investigated in (Azar and Menassa, 2012; Ioannou and Itard, 2015; Tahmasebi and Mahdavi, 2015), to name a few. Studies of the sensitivity of occupancy prediction models to their input occupancy data, however, are missing.

In the case of occupancy input data, occupants' presence information (i.e., occupancy profiles) significantly contributes to providing reliable occupancy information. Deterministic and probabilistic occupancy profiles are widely used to model occupancy information. In spite of the simplicity of deterministic models, all days of a week have been assumed to have the same profile throughout the year in these models, which results in the same level of energy consumption in spaces with similar size and other energy-related characteristics within the building (Davis and Nutter, 2010). Moreover, deterministic schedules fail to consider the variations of the energy consumption in the cases of special events. Also, the peak load of spaces may be overestimated, as these schedules consider the maximum occupancy in all spaces at the same time. However, this situation rarely happens in buildings. Thus, more precise and detailed occupancy models (i.e., probabilistic occupancy prediction models) should be integrated with simulation tools to more realistically estimate the energy consumption of buildings. Using probabilistic prediction models offers a more accurate representation of building occupancy information and helps reduce the gap between simulation results and the actual energy consumption of buildings. Occupancy prediction models use real data pertinent to the occupants' location, movement, and actions to predict the probability of an event (i.e., an occupant being present in a space) or an activity (e.g., window opening behavior) and generate the probabilistic profiles (Virote and Neves-Silva, 2012; Wei et al., 2018). Accordingly, the first step towards having reliable occupancy information is collecting the proper amount of data for a reasonable period and with an acceptable resolution level. Modern buildings are equipped with various types of sensors to collect the required data for their operation. Thus, enough amount of input data can be provided to occupancy prediction models, and eventually simulation tools, for predicting building energy performance (Moreno et al., 2016).

On the other hand, the massive amount of collected raw data requires significant data processing to provide reliable results (Khan and Hornbæk, 2011). Therefore, finding a balance between the amount of collected data and the required accuracy is important. To this end, the resolution level of collected raw data plays a key role. Different data collection periods along with various time-steps for recording and analyzing the occupancy data have been reported as shown in Table 5-1.

The resolution levels of five and 60 minutes are used more frequently to collect occupancy data. Moreover, 60 minutes is the most frequently used time-step to analyze the collected data (Liang et al., 2016). However, this time-step may not be suitable for all types of applications. In order to enhance the quality of the results generated by the probabilistic prediction models, it is vital to investigate the effect of the changes in the input parameters on the outcomes of the prediction models. This can be done through the application of sensitivity analysis. Although many research studies emphasized the importance of incorporating occupancy information in energy assessment of buildings through the usage of occupancy data (Azar and Menassa, 2011; Oldewurtel et al., 2013; Duarte et al., 2015; Imanishi et al., 2015; Kim and Srebric, 2015; Sangogboye et al., 2018), none of them investigated the sensitivity of these models to the data collection period and the resolution level used for analyzing the collected data. To address this gap, the current study aims to answer the following questions: (1) How to select the near-optimum data collection period length that results in accurate occupancy prediction? (2) How to sample the training months out of the whole period of data collection? (3) What is the effect of different temporal resolution levels used to analyze the collected data on the accuracy of the occupancy prediction model? Answering the above questions leads to selecting the most effective data collection period and resolution level, which helps the prediction model to produce reliable occupancy information. The focus of this study is on shared office spaces with multiple zones. Having different occupants with dissimilar work habits and schedules highlights the need for probabilistic prediction models to estimate the occupants' profiles. To differentiate between different occupants' profiles, zoning is required to assign a typical zone-level occupancy profile to each zone.

5.2 Methodology

5.2.1 Framework for the Proposed Sensitivity Analyses

The methodology to apply sensitivity analysis unfolds over three steps as demonstrated in Figure 5-1.

Step 1: Data collection and preparation. Considering that the energy performance of an office is mainly influenced by its occupants, the office occupants' schedules and habits determine the input to the occupancy prediction model. To this end, three tasks are fulfilled aiming to collect occupancy data and prepare the collected data for further analysis. Real-time location systems (RTLSs) can be utilized to get occupants' locations and their presence time. The occupancy data should be collected over a reasonable period using a high-resolution level (e.g., each minute) for facilitating the sensitivity analysis. Furthermore, very high-resolution occupancy data are required for some specific applications, such as security and emergency situations. In order to improve the quality of the collected data, data cleansing is applied, which comprises filling the missing data, removing duplications, and detecting outliers in the collected raw data. After cleansing the data, data processing is performed to obtain the parameters required for developing the occupancy prediction model. The spatiotemporal patterns of the occupant's behavior along with their number are derived from the collected data during this step. Since the very high resolution is not required for building energy management, the occupancy data are generated with the time intervals of one minute during the data processing phase.

Table 5-1 Implementation settings of different research studies

Reference	Data collection period	Data collection time-step (min)	Analysis time-step (min)
Balaji et al. (2013)	3 weeks	15	60
Yang and Becerik-Gerber (2014)	6 months	3	15
Chen and Ahn (2014)	20 days	5	20 & 30
Dobbs and Hencey (2014b)	58 days	1 sec.	60
D'Oca and Hong (2015)	2 years	10	60
Wang and Ding (2015)	1 week, working time	10	10
Liang et al. (2016)	1 year	5	60
Wang and Shao (2017b)	30 days	35 sec.	60
Wang et al. (2017)	7 days	1	20
Newsham et al. (2017)	31 days	15 sec.	15 sec.
Capozzoli et la. (2017)	4 months	15	15
Peng et al. (2017)	7 months	1	1
Wang et al. (2017)	10 days	5sec.	5
Jiefan et al. (2018)	22 months	60	60
Pang et al. (2018)	1 year	60	60
Zou et al. (2018)	2 days	1 sec.	30 sec.
Howard et al. (2019)	59 days	60	60
Deng and Chen (2019)	54 days	5	5
Piselli & Pisello (2019)	1 year	5	60
Kim et al. (2019)	4 months	20 sec.	60

Step 2: Occupancy prediction model. After preparing the raw data, a probabilistic occupancy prediction model is developed in this step. A variety of occupancy prediction models have been developed to generate probabilistic occupancy profiles as mentioned in Chapter 2. In this study, the prediction model developed in Chapter 4 is used. The output of the proposed inhomogeneous Markov chain model is the probabilistic profiles of each specific occupant for each day of the week. The work state of each occupant, his/her location (i.e., zone) and the total number of present

occupants can be derived from these profiles at each time-step. It is assumed that the occupants are assigned to the office with assigned seats. This assumption helps the prediction model to learn the occupants' working habits and schedules. Moreover, the proposed model is a generic prediction model that is independent of the type of office (e.g., single- and multi-occupied offices).

Step 3: Sensitivity analyses. There are two critical parameters affecting the performance of the occupancy prediction model including the data collection period for training the model and the time-step used for predicting future occupancy. In order to investigate the effects of these parameters on the accuracy of the prediction model, two sets of sensitivity analyses are applied as explained in the following sections.

5.2.2 Sensitivity Analysis of the Data Collection Period

A prediction model trained over a short data collection period cannot capture variations in occupancy profiles. Presumably, more data would result in a more reliable and accurate prediction model. On the other hand, collecting data over a long period requires more computation time and costs more to process and analyze the data. Hence, finding the optimal data collection period to balance between accuracy and computation effort is important when developing a prediction model.

Thus, the first sensitivity analysis investigates the effect of the data collection period on the accuracy of the occupancy prediction model. To do so, two steps are applied including the ranking of the months in the training dataset and then varying the length of training data used in the analysis. The latter is performed by reducing the number of training months from the total number of months in the training dataset (M months) to one month after ranking them.

The training months are sorted based on their rankings, which are defined according to the following criteria: data spread feature showing the variations in the collected data, the reliability, and the similarity between collected time-series. Figure 5-2 illustrates this procedure. The ranking is firstly done by recognizing the months (h months) with the highest and lowest variations in terms of PT and the number of occupants. This is done to accommodate the maximum and minimum possibilities regarding occupancy patterns. Figure 5-3 shows examples of the monthly average number of occupants during months with different rankings. Moreover, months with unreliable data, such as the one illustrated in Figure 5-3 that shows occupancy only for half an hour throughout the whole month, are given the lowest ranks (l months).

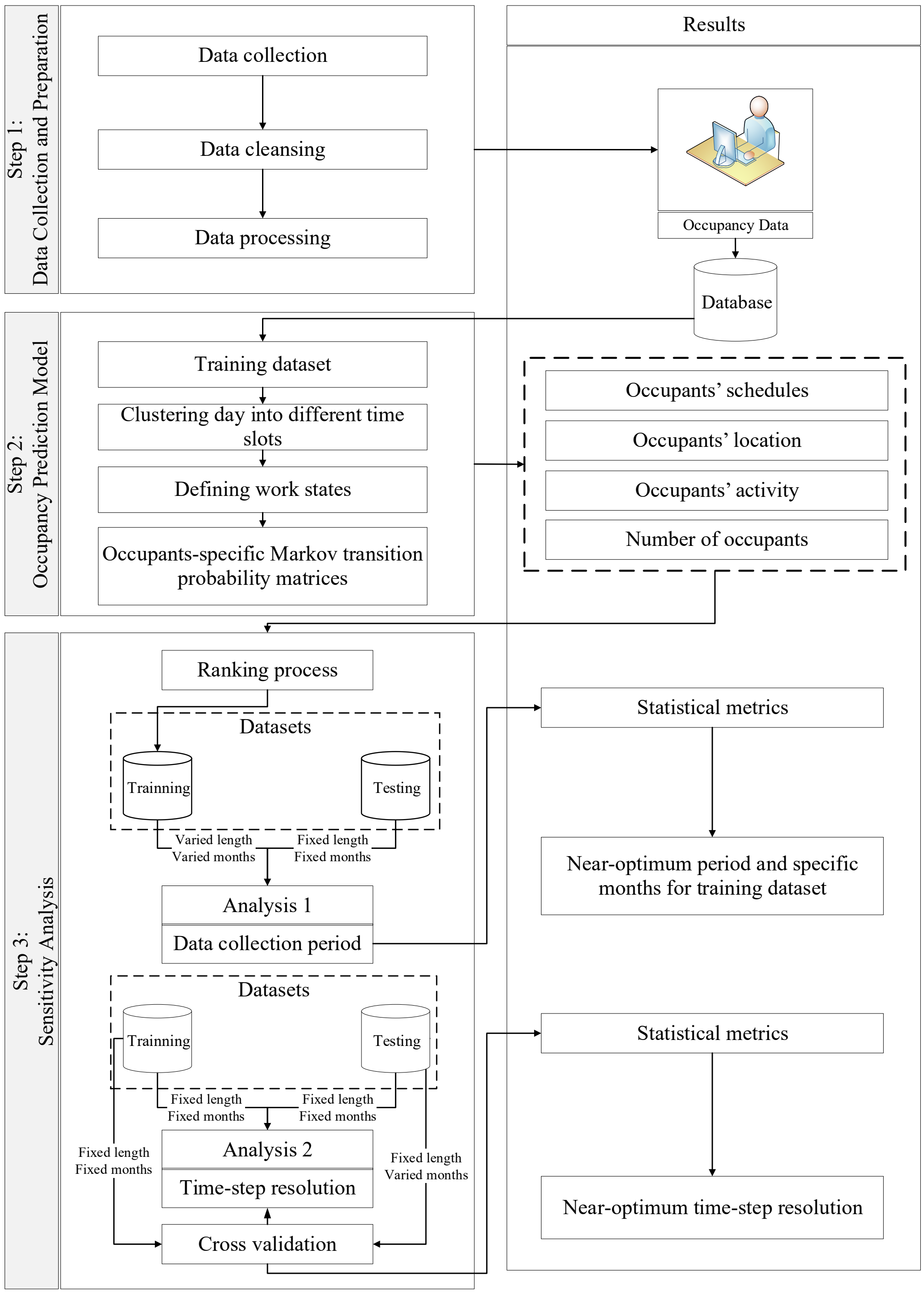


Figure 5-1 Framework of investigating the sensitivity of the occupancy prediction model

After allocating the highest and lowest ranks to h and l months, respectively, the similarity aspect of the time-series is investigated to rank the remaining months. The Kolmogorov-Smirnov (*KS*) test is a widely used test to check the similarity of time-series (Massey Jr, 1951; Justel et al., 1997; Lee et al., 2016). *KS* test is a non-parametric statistical hypothesis testing method, which can determine whether two time-series of a random variable (x) from any arbitrary distributions are similar or not (Lee et al., 2016; Christ et al., 2016). After normalizing the two time-series (p and q), as shown in Equation (5-1), the CDF ($cdf_{p'}(x)$ and $cdf_{q'}(x)$) of the two normalized time-series are derived. Then, the maximum distance between these two CDFs is calculates as the *KS* statistic (D_{stat}) (Massey Jr, 1951):

$$p_i' = \frac{p_i}{\sum_{i=1}^{T} p_i} \qquad q_i' = \frac{q_i}{\sum_{i=1}^{T} q_i} \tag{5-1}$$

$$D_{stat} = max_x \left| cdf_{p'}(x) \; - cdf_{q'}(x) \right| \tag{5-2}$$

where T is the total number of time-steps in a day. The D_{stat} shows the similarity between two time-series (i.e., the lower values of D_{stat} indicate more similarity).

In the next step, a $n \times n$ statistic matrix containing the D_{stat} for each pair of time-series of the monthly average number of occupants for the remaining n months is created. n is determined by deducting h and l from the total number of months in the training dataset (M months). In order to find the first v months with the highest similarity, all combinations of v months out of n months, $com_v^n = \binom{n}{v}$, are created. Having the D_{stat} for the pairs within each combination (i.e., $com_2^v = \binom{v}{2}$), the months in the combination with the smallest summation of $\sum_{i=1}^{com_2^v} D_{stat,i}$ are granted the next highest rank. Moreover, an average time-series (ave_v) of the number of occupants of these v months is calculated. The remaining months are $I = M - (h + l + v)$. Afterward, the *KS* test is applied on each pair of i, ranges from 1 to I months, and ave_v time-series to find the most similar month to ave_v. After determining the month with the lowest dissimilarity with respect to ave_v, the next rank is assigned to this month and the ave_v is recalculated after adding this month to the pool of v months ($v = v + 1$). These steps are repeated until all months are ranked.

After ranking the training months, the second step of the first sensitivity analysis is applied to find out the near-optimum length of the training dataset ($L_{training}$). In this regard, the length of training months is varied from M months to one month. The reduction process is done based on the rankings of the training months in which the months with lowest ranks are deducted first. At each step of the reduction process, the prediction model is trained using the selected months and the accuracy of the model is calculated against the results obtained from a testing dataset with a fixed length ($L_{testing}$) and fixed months.

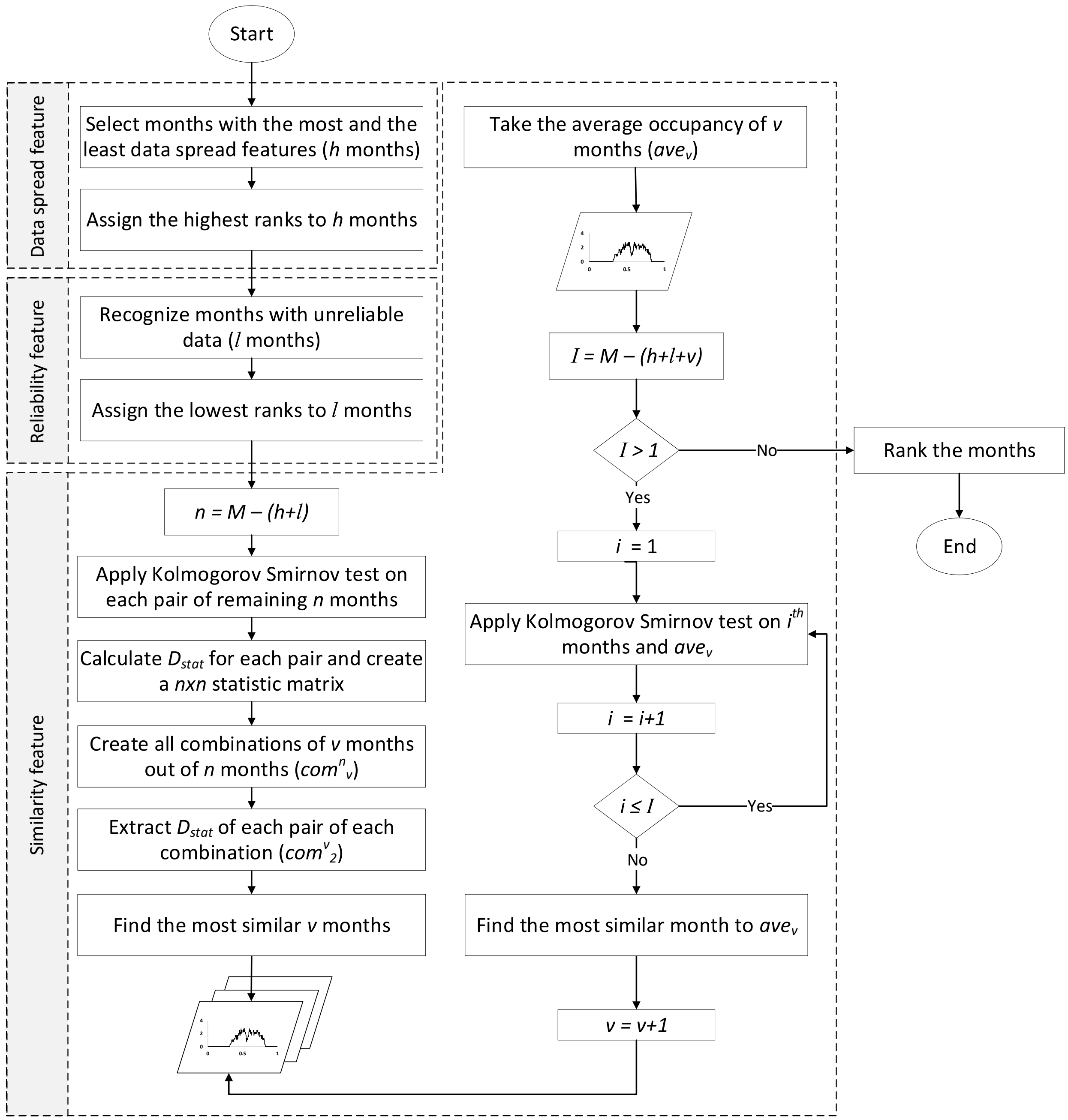


Figure 5-2 Training dataset ranking procedure

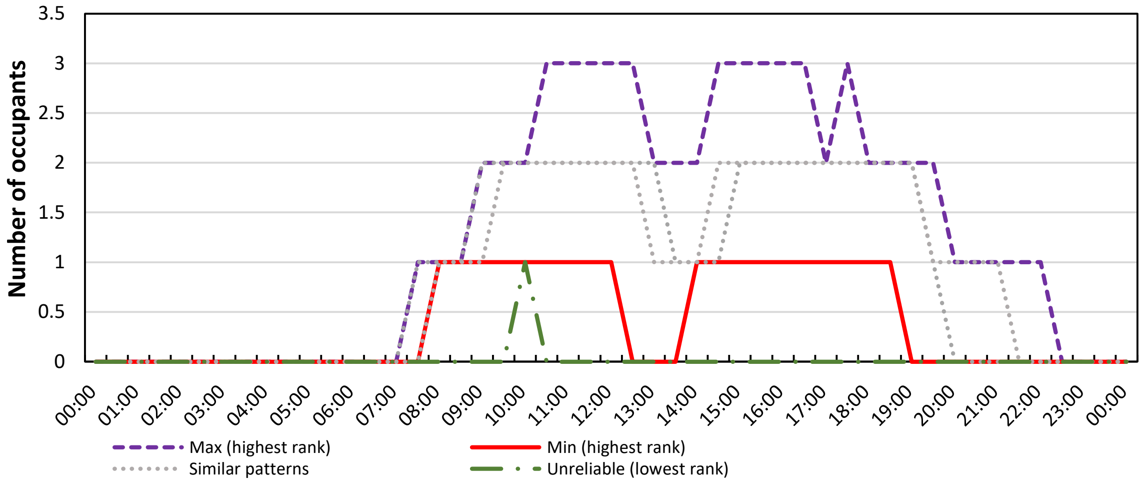


Figure 5-3 Example of monthly average occupancy patterns from months with different ranking criteria

5.2.3 Sensitivity Analysis of the Simulation Time Step

In the second sensitivity analysis, the variation in the time-step resolution used for predicting the future occupancy is explored. There are two rounds of sensitivity analyses applied for this purpose. In the first round, the optimum number of months of the training data, as found in Section 5.2.2, and a testing dataset with fixed length and a fixed number of months are used. The testing dataset is the same as the one used in the first sensitivity analysis. The time-step is changed reflecting different resolution levels and the statistical performance metrics, as explained in Section 5.2.4, are calculated.

Since the ${P_{ij}}_{o}^{d}(\mathrm{t})$ matrices are derived based on a one-minute time-step, some adjustments and updates should be applied in the case of using occupancy prediction for time-steps bigger than one minute as demonstrated in Figure 4-17. In these cases (e.g., having the time-step of n minutes), the initial state of occupancy is determined using the real-time collected data and the prediction model is run for n times to produce the occupancy patterns for the next n minutes. Then, the actual occupancy state is read again from the occupancy sensors to restart the prediction procedure for the next time-step. This update improves the accuracy of the prediction model by avoiding the accumulation of errors happening after each n-minute time-step. The prediction process is then repeated for the next n minutes and this loop is continued till the end time of the simulation.

The second round of simulation time-step sensitivity analysis validates the outcomes obtained from the first round using a cross-validation process. The ultimate result of this analysis is the optimum temporal resolution level that leads to a tradeoff between the level of accuracy in predicting occupancy and the computation effort.

For cross-validation, the complete set of collected data is partitioned into two subsets, training and testing datasets using a one-month time interval. The optimum number of months for the training dataset ($L_{training}$) is determined by means of the first sensitivity analysis as explained in Section 5.2.2. The testing dataset comprises of $L_{testing}$ months (e.g., six months) is randomly selected from the remaining collected data in several iterations. The schematic procedure of the cross-validation process is shown in Figure 5-4. In each iteration, different accuracy metrics, as explained in Section 5.2.4, are calculated.

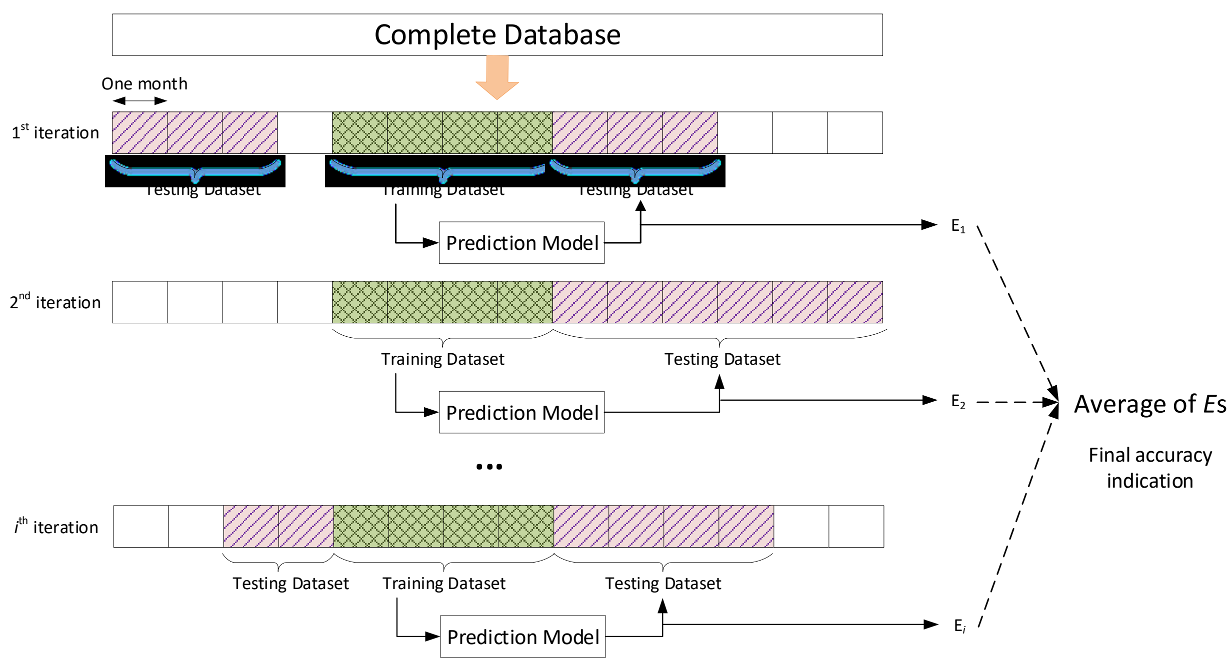


Figure 5-4 Schematic procedure of the cross-validation process

5.2.4 Statistical Metrics

In this study, R^2 (coefficient of determination) and nRMSE (normalized Root Mean Square Error) are used as the statistical performance metrics. R^2 provides a measure of how well the collected data are replicated by the prediction model. The equation of this metric is:

$$R^2 = 1 - \frac{\sum_{n=1}^{N}(L_n - P_n)^2}{\sum_{n=1}^{N}(L_n - \bar{L})^2} \tag{5-3}$$

In this study, L_n and P_n represent the real number of present occupants and prediction values, respectively. $\bar{L}$ is the mean of the real number of present occupants. N is the size of the dataset, which changes according to the selected time-step. For instance, if real data are analyzed every minute, N is equal to 1,440 for each day.

nRMSE, also known as CVRMSE (Coefficient of Variation of the Root Mean Square Error) (ASHRAE, 2002), is the RMSE divided by the mean of the data and it quantifies the size of the error relative to the mean (Granderson and Price, 2014). The equation for this metric is:

$$nRMSE = \frac{\sqrt{\sum_{n=1}^{N}(L_n - P_n)^2}}{\bar{L}} \tag{5-4}$$

A third accuracy indicator of the model is determined by averaging the obtained metrics in all iterations. Assuming E as the performance metric, the final indication after I iterations is calculated as follows:

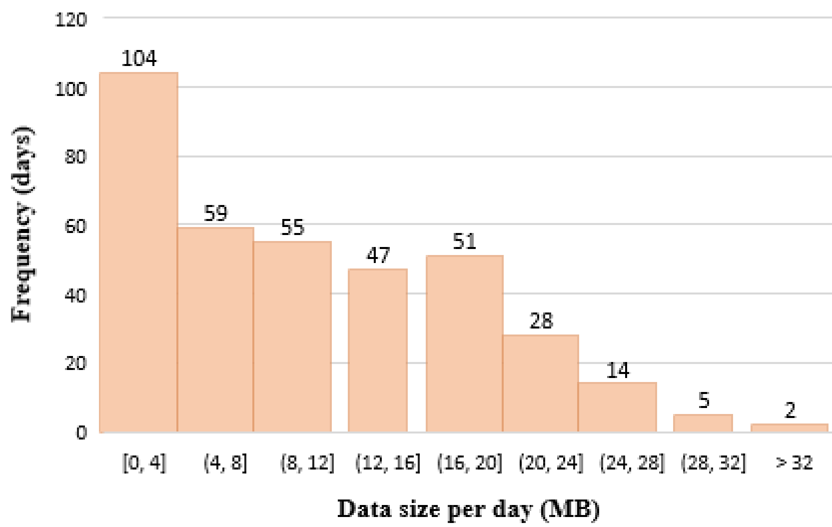
$$\bar{E} = \frac{\sum_{i=1}^{I} E_i}{I} \tag{5-5}$$

5.3 Case study and Results

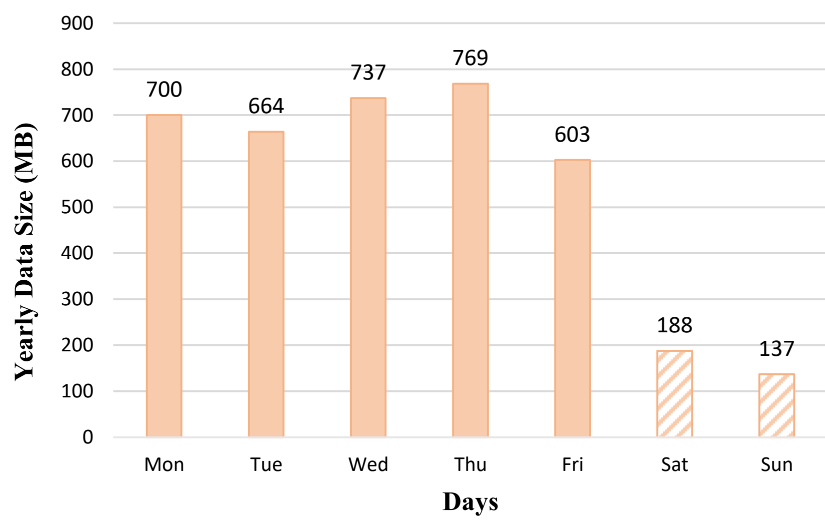
5.3.1 Data Collection and Preparation

The same case study, used in Chapter 4, is considered to demonstrate the feasibility of the proposed method in finding the suitable data collection period and resolution level to be used in the occupancy prediction model.

The occupancy data are collected for 18 months using a very high temporal resolution level (i.e., each second). Collecting the occupancy data with the high resolution of one second generated about 3.8 GB of yearly raw data in total resulting in more than 14 million data points. Figure 5-5 shows the distribution of the size of the collected data over one year of data collection. The distribution is a right-skewed distribution, which indicates that most of the collected data per day (104 days) has a size smaller than 4 MB. Moreover, the daily distribution of the collected data size shows almost equal data size for weekdays with the highest size for Thursdays and the lowest for Fridays during weekdays. The results indicate that the occupants spend more time in the office on Thursdays (most probably due to preparation for Fridays' meetings or the end of the week deadlines) and leave the office earlier on Fridays.



(a) Yearly distribution



(b) Daily distribution

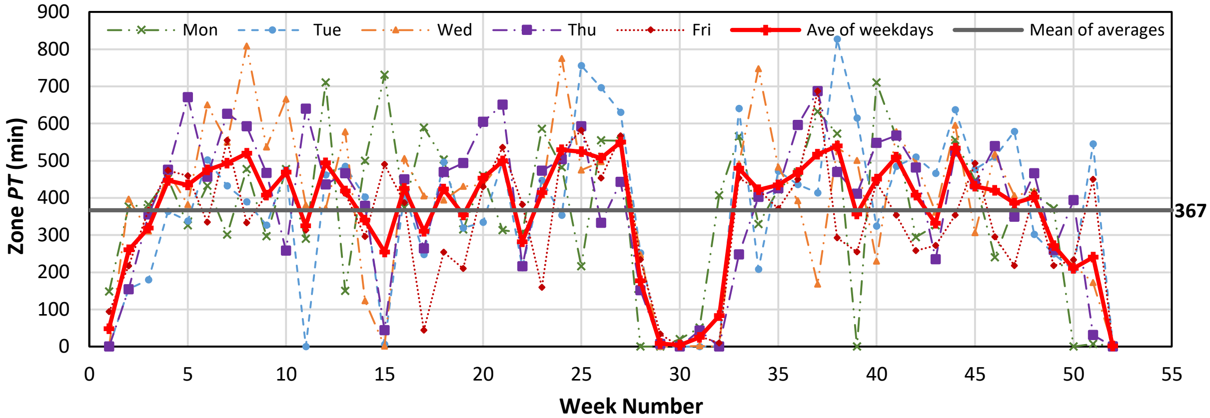
Figure 5-5 Distribution of the collected data size

Temporal variations in occupancy presence data are shown by analyzing the changes in the occupancy PT for two-time horizons (i.e., daily and monthly) based on the one-year collected data. Using the start and end times of occupancy, the total PT of occupants for each day of a week and at each zone are derived. The raw data characteristics are shown in Table 5-2 by measuring the central tendency and dispersion of daily PT at each zone. According to this table, the mean PT increases by 20% and 15% when excluding weekends with lower presence time at Zones 1 and 2, respectively. Although 50% of days (excluding weekends) have PT above 406 minutes for Zone 1, PT of zero occurs most often compared to other values. High values of the standard deviation show that the PT at both zones is spread out over a wide range of values. On the other hand, less variability is observed around the mean of the PT based on the standard errors' values.

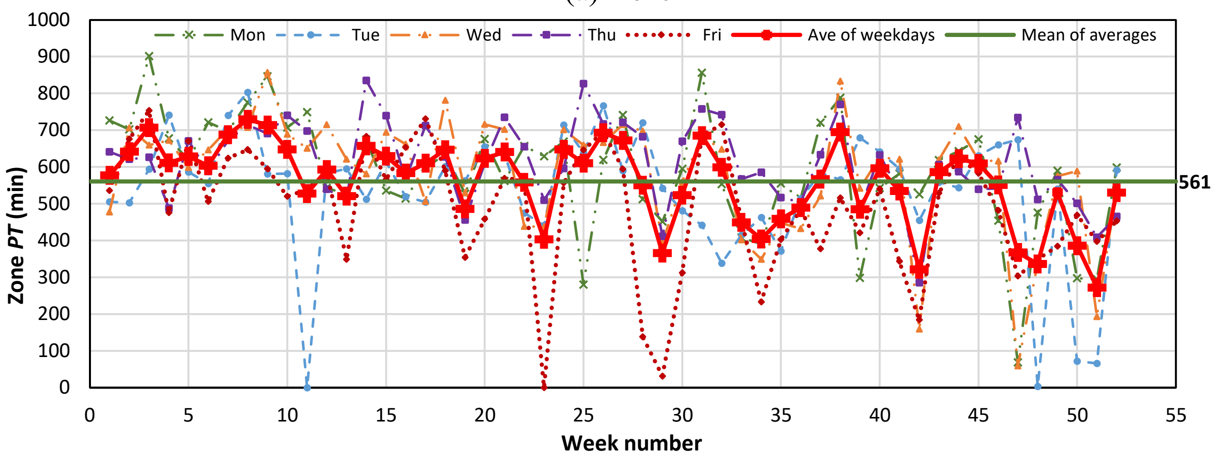
Table 5-2 Descriptive statistics of collected raw data

		Daily PT at Zone 1 (min) Whole year		Daily PT at Zone 2 (min) Whole year	
		Including weekends	Excluding weekends	Including weekends	Excluding weekends
Measure of central tendency	Mean	292	367	478	561
	Median	306	406	530	586
	Mode	0	0	0	595
Measure of variability	Standard deviation	221	198	225	162
	Standard error	12	10	12	8
	Min	0	0	0	0
	Max	827	827	901	901

Figure 5-6 depicts the total PT at each zone for weekdays. The beginning of the year shows shorter PT due to the holidays. For the first half of the year (i.e., the first 27 weeks), the average PT is around 400 minutes for Zone 1. This number drops almost 17% to 333 minutes for the second half of the year. However, the average PT of the first half of the year rarely goes under 300 minutes. The same trend is observable for Zone 2. This indicates that office occupants tend to spend more time outside of the office when the outside weather gets better during the spring and summer. Therefore, it is important to consider seasonal variations when defining occupancy profiles. The same data are generated for weekends for both zones, which are not included here due to space limitations. As expected, the occupancy rate of the office is much lower during the weekends. However, it is important to know that the office is occupied for several hours during most weekends.



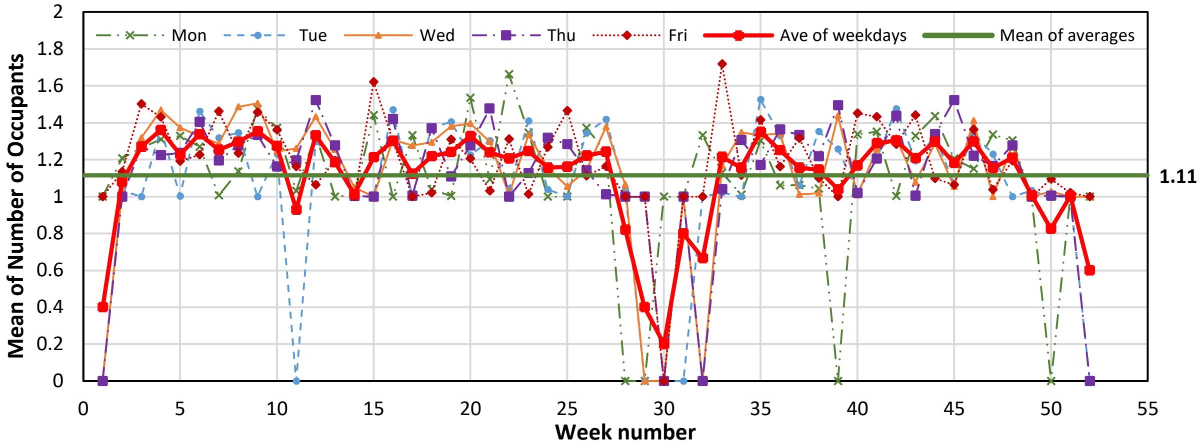
(a) Zone 1



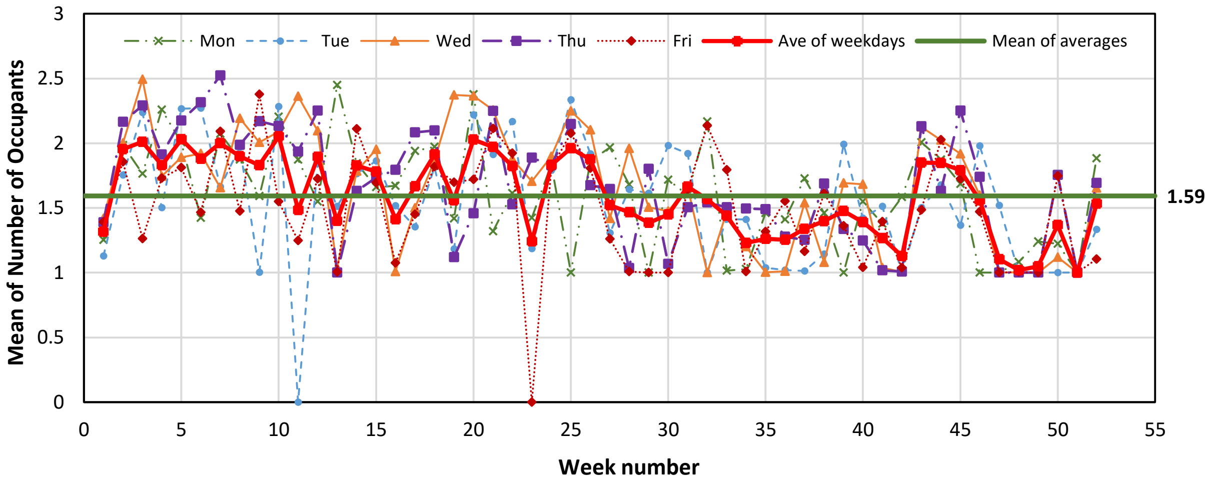
(b) Zone 2

Figure 5-6 Variation in zone occupancy ***PT*** for different days of a week and the average ***PT***

Variations in the number of present occupants at each zone are also investigated in Figure 5-7. The expected values of the mean of average occupancy PT ($\mu^{z}_{PT_{ave}}$) and the number of occupants at each zone ($\mu^{z}_{O_{ave}}$) during a year are also calculated, which are shown by horizontal lines in Figure 5-6 and Figure 5-7 (a) and (b). The lower number of the mean of average present occupants (i.e., 1.11 and 1.59 in Zones 1 and 2, respectively) compared to the peak occupancy (i.e., 3) demonstrates that using the maximum number of occupants all the time (according to the common practice) results in overheating or overcooling space. Moreover, wasting energy would be the result of conditioning zones with no occupancy. In addition, Zone 2 shows 43% higher average the number of occupants compared to Zone 1. This indicates the importance of considering the occupancy data at zone level instead of room level. It is especially important when applying local control strategies to building energy-consuming systems (Liu et al., 2016). Local control strategies exploit the spatiotemporal variations of the zone-level occupancy patterns (including occupants' locations and occupancy duration) to adjust building systems.



(a) Zone 1



(b) Zone 2

Figure 5-7 Variation in the number of occupants for different days of a week and the average occupancy

The occupancy daily *PT* over different months of the year is illustrated in boxplots for both zones as demonstrated in Figure 5-8. Boxplots summarize sets of data by showing the shape of the data distribution, their central value, and variability. Therefore, they are one of the best ways to compare different data sets. The box, also known as the interquartile range (IQR) represents the middle 50% of data and the top and bottom whiskers extend to the highest and lowest values in a data set. One-quarter of the data set span in the lower whisker and the highest 25% of values fall within the upper quartile (Massart and Smeyers-verbeke, 2005; Wickham and Stryjewski, 2011).

It can be seen from this figure that the *PT* of occupants in Zone 2 on average exceeds that of Zone 1 during all months of the year. The highest difference happens during the month of January in which all days at Zone 1 have lower *PT* than the *PT* of the 75% of the days at Zone 2. Almost the same pattern is detectable for the next three months (i.e., February, March, and April). These boxplots also show that the PT is more consistent in Zone 2 and remains at upper levels while Zone 1 PT is more variable, especially at lower levels. In Zone 2, the PT is higher than six hours for 75% of the days of each month except for December. The first half of the year shows the most symmetric data for Zone 2. Zone 2 consistent PT makes predictions more dependable than the highly variable PT of Zone 1. The collected data are skewed to lower values for most months of the year at Zone 1, which means that most days have relatively short PT. This observation is aligned with the conclusion that occupancy in Zone 2 is much higher than that of Zone 1.

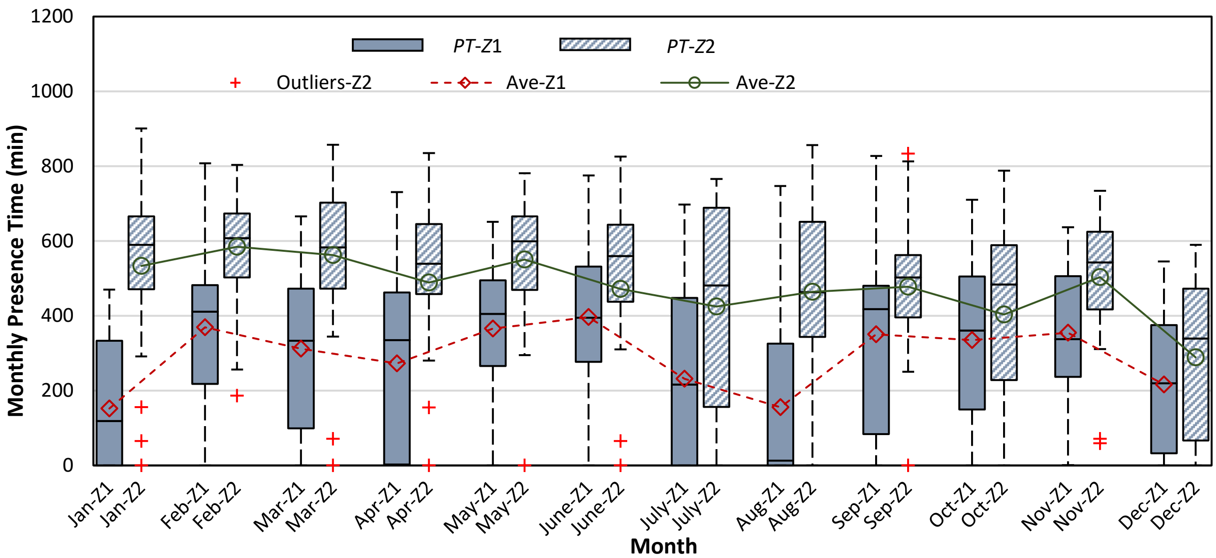


Figure 5-8 Variation in occupancy $\boldsymbol{PT}$ for different months of a year

Variations in the median of PT can also be observed in this graph. January shows the lowest occupancy for Zone 1 followed by December. The median of PT of Zone 1 then starts to increase and is almost stable during the next five months. Furthermore, distinguished drops can be detected during July and August. There is an increase in the PT for the following months. However, the median shows lower values for the second half of the year compared to the first half at both zones. These drops and variations could be due to the spring and summer breaks. This is an indication of seasonal variations in the PT. December also has the lowest occupancy record for Zone 2. The much longer whiskers for PT of Zone 2 in August indicate that PT varies more widely during this month. The peak of the median occupancy is observed during June and February for Zones 1 and 2, respectively. Some outliers are observed at Zone 2. These abnormal values can affect the overall observation due to their very high or low extreme values, and hence should be discarded from the data series. Removing these points avoid underestimating or overestimating the PT.

Overall, the above discussion shows the importance of applying the effect of the temporal behavior of occupants when predicting their presence patterns.

5.3.2 Occupancy Prediction Model

To estimate the occupancy profiles using the probabilistic inhomogeneous Markov chain occupancy prediction model, all states are labeled to show the transition probabilities from one state to another according to Table 4-1.

As discussed in Section 5.2.1, a one-minute prediction time-step is considered to predict the office occupancy patterns. Having the occupants' zones for every minute time interval, the distribution of the time being spent in the office's zones and outside is determined for each day of a week (i.e., $ns_{o,i}^{t,d}$). After calculating the number of transition occurrences (i.e., $ntr_{o,ij}^{t,d}$), the transition matrices corresponding to each time-step of each day of a week would be calculated using the average values of $tr_{o,ij}^{t,d}$ and $s_{o,i}^{t,d}$ for that specific day of the week throughout the whole year using Equations (4-8) and (4-9). For instance, the transition matrices of occupant o_5 on Wednesdays at 01:04 pm and 02:14 pm are shown below:

$$
\begin{gathered}
P_{ij}{}_{o_5}^{Wed}(01{:}04\text{ pm}) = \begin{bmatrix} 0.7934 & 0.0147 & 0.0295 & 0.0147 & 0.1475 \\ 0.0147 & 0.9796 & 0.0008 & 0.0004 & 0.0004 \\ 0.0295 & 0.0008 & 0.9600 & 0.0008 & 0.0086 \\ 0.0147 & 0.0004 & 0.0008 & 0.9796 & 0.0004 \\ 0.1475 & 0.0004 & 0.0086 & 0.0004 & 0.8350 \end{bmatrix} \\
P_{ij}{}_{o_5}^{Wed}(02{:}14\text{ pm}) = \begin{bmatrix} 0.7504 & 0 & 0.1193 & 0.0217 & 0.1085 \\ 0 & 0 & 0 & 0 & 0 \\ 0.1193 & 0 & 0.8233 & 0.0095 & 0.0477 \\ 0.0217 & 0 & 0.0095 & 0.9600 & 0.0086 \\ 0.1085 & 0 & 0.0477 & 0.0086 & 0.8350 \end{bmatrix}
\end{gathered}
\tag{5-6}
$$

5.3.3 Sensitivity Analyses

5.3.3.1 Sensitivity Analysis of the Fata Collection Period

In the first step of the sensitivity analysis 1, out of all collected data, the first 12 months are selected as the training dataset and the test dataset comprises of the remaining 6 months. Time-step resolution of one minute is selected for this analysis to provide the highest accuracy. To investigate the sensitivity of the time-dependent inhomogeneous Markov chain occupancy model to the data collection period, the number of training months is reduced from 12 months to one month according to the ranking procedure explained in Section 5.2.2. The ranking schema is shown in Table 5-3 for each zone. In this table, the lower numbers show the higher ranking and subsequently the better quality of the collected raw data. Thus, the months with lower ranks are firstly removed from the training dataset, such as the 12th and 4th months for Zone 2.

Table 5-3 Rankings of the months in the training dataset

	Ranking	1	2	3	4	5	6	7	8	9	10	11	12
Zone 1	**Month No.**	2	12	6	7	11	5	3	9	1	10	4	8
Zone 2		3	9	2	7	5	1	6	10	8	11	4	12

Considering the occupancy prediction at the zone level, the results of the first sensitivity analysis are demonstrated in Figure 5-9 for Wednesdays. The same results are obtained for other days of the week, which, for brevity, are not included in this study. The performance measurement of the prediction model using R^2 is shown in this figure. Each point shows the R^2 of the prediction model trained by a certain number of months and evaluated using the testing dataset (with a fixed number of 6 months). Overall, the performance of the prediction model in estimating the occupancy rate is higher for Zone 2 compared to Zone 1. That is mainly due to the higher quality of collected raw data in Zone 2 to that of Zone 1. Occupants in Zone 2 were more dedicated to wearing their tags all the time and spent more time in the office compared to those in Zone 1.

As shown in Figure 5-9, the values of R^2 improves by increasing the number of months of the training dataset. The R^2 reaches the highest value at three and four months of data collection as the training set in Zones 1 and 2; respectively, showing the highest accuracy in predicting the occupancy. Thus, three (i.e., months 2, 12, and 6) and four (i.e., months 3, 9, 2, and 7) are selected as the optimum numbers of months required to train the prediction model in Zones 1 and 2; respectively, and used for the second sensitivity analysis. Although there are more variations in the values of R^2 for Zone 1 after three months, the performance of the prediction model in estimating the occupancy at both zones almost reaches a plateau after the selected optimum months of data collection as the training datasets.

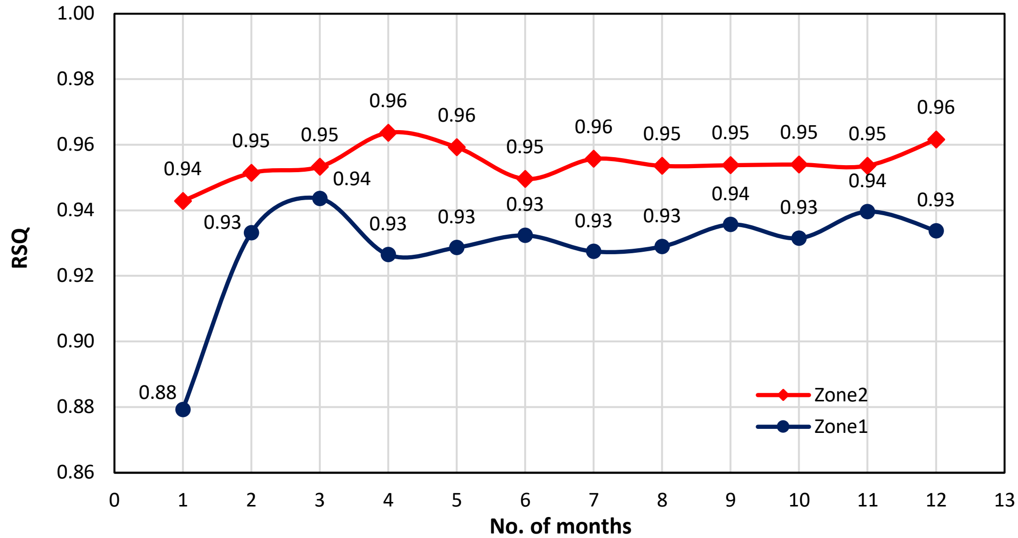


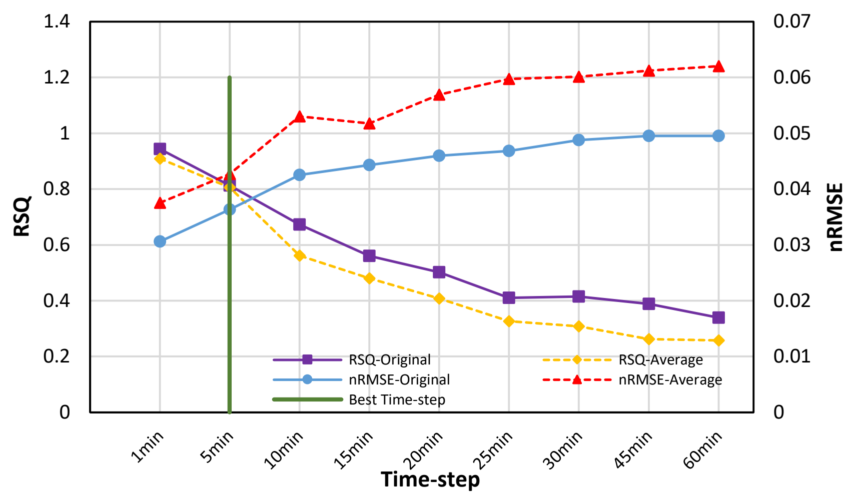
Figure 5-9 Performance comparison for different lengths of the training dataset (for Wednesdays)

5.3.3.2 Sensitivity analysis of simulation time-step

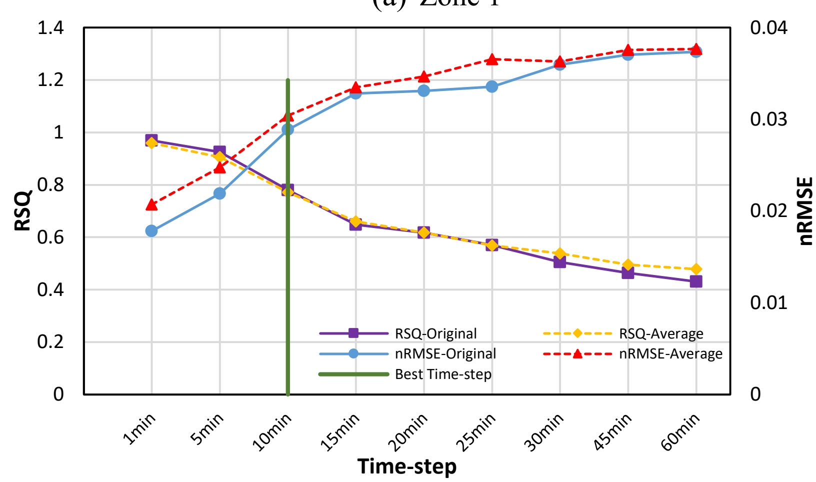
In the second analysis, the variation in the time-step resolution is explored. In this regard, after finding the optimal number of training months, different time-steps are considered to evaluate the accuracy of the occupancy prediction model. Thus, the first round of the sensitivity analysis 2 is applied on the optimum three and four months of training data, as found in the previous analysis for Zones 1 and 2; respectively, and six months of the test data. The test dataset is the same as the one used in the first sensitivity analysis. The time-step is changed reflecting eight resolution levels and the values of R^2 and nRMSE are calculated using Equations (5-4) and (5-3).

The results are demonstrated in Figure 5-10 by solid lines. As shown in this figure, five- and 10-minute time-steps result in an R^2 value of 0.8 for both zones. After these levels, the accuracy drops especially for Zone 1. The nRMSE of Zone 1 is twice as that of Zone 2 for all resolution levels, which indicates that the discrepancy between the predicted and the real collected data are occurring more frequently for Zone 1. These results prove the lower quality of the collected data at Zone 1 compared to that of Zone 2, which are in agreement with the observations of the collected raw data explained in Section 5.3.1 indicating the better quality of the collected data at Zone 2. It can be concluded that the required resolution for an acceptable accuracy level is dependent on the quality of the collected data. The lower quality of raw data leads to the need for smaller time-steps in order to have the desired performance.

In order to validate the obtained outcomes, a cross-validation process with 10 iterations is employed on the testing dataset as explained in Section 5.2.3. In this study, a set of six months comprising the test dataset is randomly selected among all months of the data collection except the months used for the training dataset at each iteration. The performance of the occupancy prediction model is then evaluated against the selected test dataset. The final performance assessment is obtained by averaging the results calculated at each iteration as shown in Equation (5-5). As illustrated by dashed lines in Figure 5-10, the results are aligned with those calculated in the first round, which shows the validity of the proposed analyses.



(a) Zone 1



(b) Zone 2

Figure 5-10 Performance assessment of the prediction model for different resolution levels

5.4 Summary and Conclusions

Occupancy-related parameters are recognized as influential parameters affecting either positively or negatively the energy operation in buildings. In order to find comprehensive probabilistic occupancy profiles, which capture the variations in occupancy presence and diverse activities of different occupants within a building, probabilistic occupancy prediction models should be leveraged. Improving the reliability of probabilistic prediction models depends on the accuracy of the input parameters used to develop these models. There are two critical parameters affecting the performance of occupancy prediction models including the length of the data collection period for training the model and the time-step used for predicting future occupancy.

In this chapter two sensitivity analyses are performed to investigate the dependencies of the outcomes of an occupancy prediction model on the changes to its input parameters. Using the proposed method in similar cases provides the following benefits: (1) ranking months of data collection using the proposed ranking procedure based on the data spread feature, the reliability, and the similarity between collected time-series; (2) determining the near-optimum length of the data collection period required at each zone of a space; (3) selecting the near-optimum training dataset with the length found in the previous step; and (4) finding the most satisfying temporal resolution level for analyzing the occupancy data assuring acceptable accuracy in occupancy prediction.

To this aim, the occupancy data are collected for 18 months in an open-plan office and different analyses are conducted to study various aspects of the occupancy data and their impact on the accuracy of the occupancy prediction model. The key insights drawn from the analyses are summarized as follows:

1) Processing the collected raw data indicates that the occupancy presence data are spread out over a wide range of values at both zones.
2) The sensitivity analysis of the data collection period results in the selection of three and four months of data collection as the optimum number of months required to train the prediction model for Zones 1 and 2, respectively.
3) Five- and 10-minute time-steps showed the acceptable value of R^2 for Zones 1 and 2, respectively.

4) Further examination of the outcomes obtained from the first round of the sensitivity analysis 2 is performed through the cross-validation process. This step validates the model performance in capturing variability in data patterns and effectively predicting occupancy profiles. The results are aligned with those calculated in the first round, which shows the validity of the proposed analyses.

It can be concluded that the required resolution for an acceptable accuracy level is dependent on the quality of the collected data. The lower quality of raw data leads to the need for smaller time-steps in order to have the desired performance.

Overall, the above results show the importance of investigating the effect of the temporal behavior of occupants when predicting their presence patterns. Obvious differences in the number of occupants and patterns in different zones indicate the importance of considering the occupancy data at zone level instead of room level. This information is especially crucial when applying local control strategies to building energy-consuming systems.

As future work, seasonal changes in occupancy patterns can be studied in detail by having different training datasets for different seasons. Other methods can be used for ranking the months in the training dataset, such as clustering. Furthermore, the results of the proposed method can be used to quantify the impact of the occupancy prediction model, tuned based on the most effective data collection period and resolution level, on the energy performance of buildings. Energy models coupled with accurate occupancy prediction model not only lead to optimizing the operation of buildings but also can improve the thermal satisfaction of occupants. Future work could also consider cost-benefit analysis to investigate the balance between the cost of using different RTLSs, collecting data over a long period, and the gains of using them in the real world. In addition, solutions to solve privacy issues when collecting detailed occupancy data should be investigated.

CHAPTER 6 SIMULATION-BASED MULTI-OBJECTIVE OPTIMIZATION

6.1 Introduction

While the effects of occupant-centered control strategies on the energy performance of office buildings have been studied by various researchers, most of the works are restricted to deterministic occupancy profiles. Although some diversity has been considered by using different deterministic schedules for workdays and weekends, all workdays are considered to have the same profile throughout the year (Davis and Nutter, 2010). This assumption disregards the important impact of the dynamic variations in occupants' profiles. Even when these variations are considered through the usage of probabilistic occupancy profiles, all days of a week are treated equally by using fixed pre-defined profiles (Yang and Becerik-Gerber, 2014; Goyal et al., 2015; Sehar et al., 2017). This makes these occupancy models not mature enough for representing real occupancy patterns. Furthermore, as mentioned before, generally there is an inverse relationship between the energy consumption of operational systems and the comfort level of occupants using these systems. Thus, the optimal operation of buildings' energy-consuming systems is of great importance for minimizing the energy consumption of the building while satisfying the occupants.

In this chapter, a new simulation-based multi-objective optimization model of the energy consumption in open-plan offices based on occupancy dynamic profiles and occupants' preferences is developed. Using a proper sensing technique to distinguish between different occupants in open-plan offices and detect occupancy patterns results in differentiating the temporal behavior of different occupants. Based on the derived occupancy information, an occupancy prediction model can be developed using different stochastic methods to predict future occupancy information as discussed in Chapter 4. Furthermore, the results of the prediction model have to be periodically corrected (e.g., every 30 minutes) based on the RTLS data. However, the full integration of the prediction model with the simulation-based optimization model is beyond the scope of this chapter. Instead, for the sake of simplicity, real occupancy data is used as a proxy of the occupancy prediction results. As such, occupant-specific dynamic profiles are developed. Having the occupancy dynamic profiles along with the indoor environmental conditions, this study contributes to the exploration of solutions produced by the integration of a simulation model with a multi-objective optimization process. This integration eventually leads to the identification and application of the optimal local control strategies of the building energy-consuming systems. The objectives of this chapter are: (1) Developing high-resolution dynamic occupancy profiles based on RTLS data to represent temporal variations of occupancy patterns in open-plan offices; (2) Integrating the energy simulation model with the optimization algorithm to optimally control the building energy-consuming systems and to analyze the trade-off between buildings' energy consumption and occupants' comfort; and (3) Developing local control algorithms for building energy-consuming systems. These objectives aim at assisting decision-makers in evaluating optimized occupancy-centered building operations. The novelty of the proposed method is also related to the consideration of full-year real occupancy profiles with a high resolution of 1-minute capturing the stochastic behavior of occupants in full. Thus, there are no typical occupancy patterns, rather each day has a unique occupancy profile, which is the result of fully exploiting occupancy data. This consideration coupled with detailed occupancy-centered building systems' operation and control are used to optimize building performance. Moreover, unlike other studies (Capozzoli et al., 2017), building systems' schedules not only work based on the start and end times of occupancy but also, they follow occupancy patterns throughout the day.

6.2 Occupancy Module

There are many types of information determining the accuracy of the dynamic occupancy profiles including the duration of the occupants' presence, their locations in different zones of a building, and their preferences. New RTLSs can provide the location and duration of presence while the preference data can be collected by a survey. The collected data are then processed using occupant data analytics (data processing) to derive the important occupancy features, such as the number of present occupants, periods of absence and presence, and other occasional variations in the occupants' profiles. The procedure used to develop dynamic occupancy profiles is illustrated in Figure 6-1. In this figure, Z shows the total number of zones in the office. Since different resolution levels are required for controlling different building systems, as will be explained in Section 6.3.4, two models with two resolution levels of i and m ($i < m$) are required to produce dynamic occupancy profiles. The first set of profiles is generated using a high-resolution level of i minute (e.g., one minute). For each day of a week (d) and for each hour (n) of the daily total occupncy time (TOT_d), the zone status ($Status_{z,i}$) at time-step i is derived from the collected data at each zone (z). In additon, the duration at which the zone is occupied (t_{occ}) is recorded and compared with the unoccipied duration (t_{uocc}). In this way, the status of the zone *z* during each m time-step ($Status_{z,m}$) will be determined based on compring t_{occ} and t_{uocc} and selecting the bigger value. The zone occupancy ratio ($OccR_{z,m}$) is calculated by dividing the summation of number of occupants who occupied the zone during each *i* minute within the *m* time-step by the total number of occupants assigned to that zone ($TOT_{z,occ}$). The derived occupancy profiles are used as a basis for the development of occupants-centric local control strategies as explained in Section 6.3.4. These profiles can distinguish between different occupants' schedules and habits by capturing their temporal behavior. Dynamic occupancy profiles not only enable the application of schedule-based local control of building systems but also, they reflect the stochastic nature of occupants' schedules in the simulation model. This would help to effectively analyze the energy performance of the building under uncertainties.

6.3 Simulation-based Multi-objective Optimization Module

As mentioned in Section 3.2, the optimal operation of building energy-consuming systems is dependent on finding a balance between their energy cost and the occupants' satisfaction. Simulation techniques can be used to investigate the effect of different control strategies on building energy consumption and the occupants' satisfaction. This is done by performing sensitivity analysis on the settings of the energy-consuming systems to find how changes in these settings affect the performance of the simulation model. However, simulation alone cannot explore the whole search space of a complex energy efficiency problem; therefore, optimization methods are required to investigate all the possible combinations of the settings. The developed optimization and simulation models are respectively explained in detail in Sections 6.3.1and 6.3.2, followed by a description of the integration of two models in Section 6.3.3. Local control strategies are also discussed in Section 6.3.4.

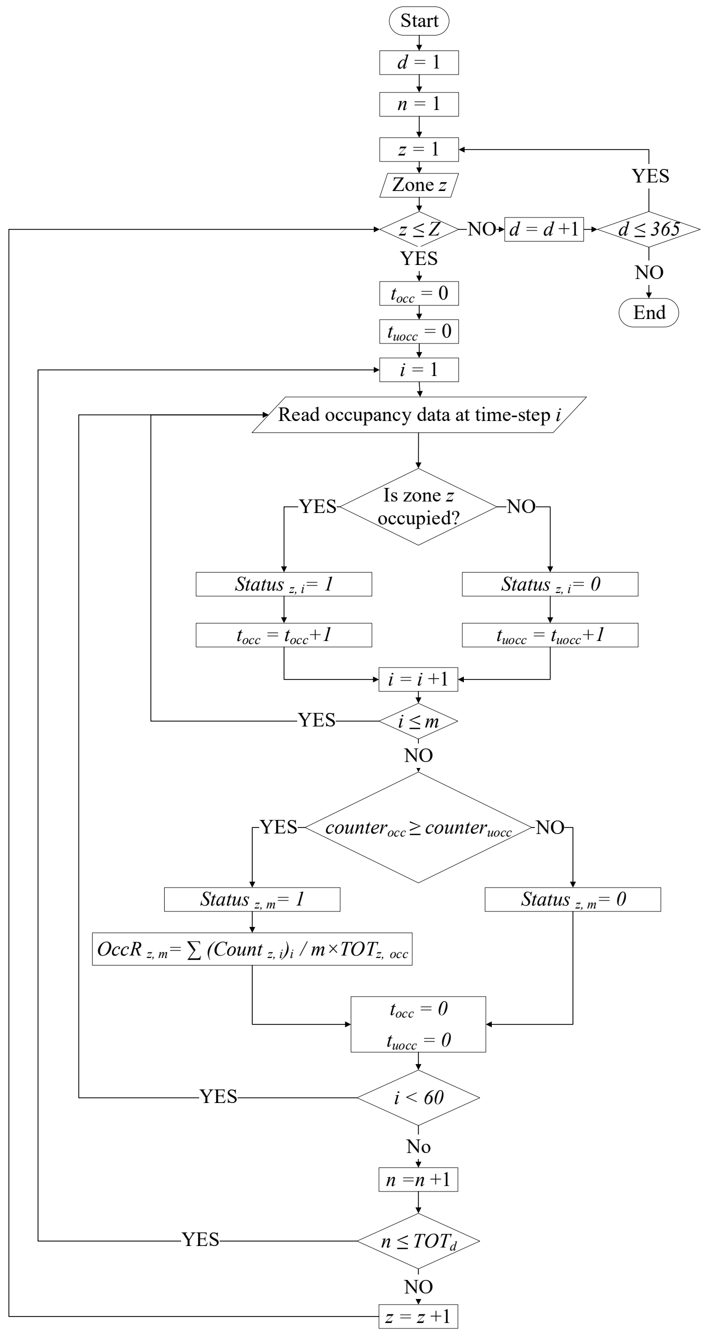


Figure 6-1 Flowchart for developing dynamic occupancy profiles based on RTLS data

6.3.1 Optimization Model

The mathematical representation of the optimization algorithm is expressed as follows:

$$\underset{X^* \in R^n}{arg\ min} \quad F(X) = [DC(X), \mathrm{E}(X)]^T \tag{6-1}$$

$$Subject\ to \qquad g_l(X)$$

where $X^* \in R^n$ is the optimal decision variables vector, and $F(X) = [DC(X), \mathrm{E}(X)]^T$ is the vector of the discomfort (DC) and energy (E) objective functions. $g_l(X)$ is the vector of the constraint functions that could take into account the regulations and the occupants' preferences regarding the HVAC and lighting systems' settings. The inputs to the optimization model are the sensor data as well as the information from the occupancy module. These inputs are the environmental conditions (i.e., the HVAC and lighting systems' settings), occupants' preferences and their dynamic profiles.

In this study, cooling and heating set-points for each zone along with the illuminance level of each zone are the optimization decision variables. In this study, minimizing discomfort hours (all clo), which is the total discomfort hours when winter or summer clothes are worn (ASHRAE 55, 2010) and minimizing energy consumption are the ultimate goals of the proposed integrated model. The level of occupants' satisfaction regarding the thermal environmental conditions defines the thermal comfort of occupants. Several mathematical models have been proposed exploring the correlation between thermal comfort variables to predict the thermal satisfaction of occupants. Among these models, the Graphic Comfort Zone method suggested by ASHRAE Standard 55-2010 (ASHRAE 55, 2010) is used in this study to measure the discomfort hours. Based on this method, the total number of discomfort hours is calculated based on whether the humidity ratio and the operative temperature are within the regions provided in ASHRAE Standard 55-2010. These regions are derived from the Predicted Mean Vote (PMV) and Predicted Percent Dissatisfied (PPD) indices developed by Fanger (1972). According to ASHRAE 55, the PMV index between +0.5 and -0.5 can be used as an indication of the thermally comfortable environment when setting the zone cooling and heating temperatures (ASHRAE 55, 2010; Sehar et al., 2017). Hence, the optimal control strategies override the normal heating and cooling set-point temperatures at zone level to maintain the PMV within the comfort range. Therefore, unlike global set-point adjustment, occupants' thermal requirements can be met at zone level resulting in fewer discomfort hours.

The energy (E) objective function is defined as a combination of the energy consumption of the HVAC system (E_H), which is the summation of cooling ($CE^s_{t,z}$) and heating ($HE^s_{t,z}$) power consumption, and lighting (E_L) system as shown below:

$$E(X) = Min\big(E_H(X) + E_L(X)\big) \tag{6-2}$$

$$E_H(X) = \sum_{s=1}^{S}\sum_{z=1}^{Z}\sum_{t=1}^{TOT}(CE^s_{t,z} + HE^s_{t,z}) \tag{6-3}$$

$$E_L(X) = \sum_{s=1}^{S}\sum_{z=1}^{Z}\sum_{t=1}^{TOT} LP^s_{t\,,z} \tag{6-4}$$

The HVAC system heating/cooling load is dependent on the internal heat gains (IHG) including gains from occupants (IHG_{occ}), lighting (IHG_l), and equipment (IHG_{eq}). The cooling energy at time-step *t* at zone *z* during season *s* ($CE^s_{t,z}$) is calculated using Equation (6-5) (EnergyPlus, 2015), where COP is the coefficient of performance of the system. Focusing on occupancy, the summation of the latent (QL_{occ}) and sensible heat gains (QS_{occ}) comprises the total occupancy heat gain and are calculated as follows (Thomas, 2018):

$$CE^s_{t,z} = {}^{cooling\ loads}\!/_{COP} \tag{6-5}$$

$$QL_{occ} = Count_z \times LHG_{occ} \tag{6-6}$$

$$QS_{occ} = Count_z \times SHG_{occ} \times CLF \tag{6-7}$$

Where $Count_z$ is the number of occupants at each time-step t at zone z. LHG_{occ} and SHG_{occ} are the latent and sensible heat gain per person for the type of activity performed in the zone, respectively, and are derived from standards. Since part of the sensible heat generated by occupants is absorbed by the surroundings and then gradually released into the zone, a cooling load factor (CLF) is considered when calculating QS_{occ} to reflect this time delay. This factor is also obtained from standards (e.g., ASHRAE). This factor is not needed for LHG_{occ}, which is instantaneously added to the zone (EnergyPlus, 2015). The same concepts are used to calculate the heating energy.

The lighting local control strategies are applied in near-real-time using 1-minute occupancy data. Two different resolution levels of 30- and 60-minutes are used to control the HVAC system, as explained in Section 6.3.4. The optimization results are shown as the summation of zone energy consumption and the number of discomfort hours in a year using Equations (6-3) and (6-4). This is done to have an overall estimation of the effect of optimal local control strategies on building energy performance and comfort.

The lighting power is calculated using Equation (6-8) (EnergyPlus, 2015):

$$LP^{s}_{t,z} = \frac{LE_z \times I^{s}_{t,z} \times A_z}{100} \tag{6-8}$$

where:

$LP^{s}_{t,z}$: Lighting power at time-step t at zone z during season s (W)
LE_z: Lighting energy at zone z (W/m^2/100 lux)
$I^{s}_{t,z}$: Zone illuminance level at time-step t at zone z during season s (lux)
A_z: Zone floor area (m^2)

The discomfort (DC) objective function is defined as the normalized summation of all discomfort time at all zones:

$$DC(X) = \frac{\sum_{z=1}^{Z}(A_z \sum_{s=1}^{S} \sum_{t=1}^{TOT} DT^{s}_{t,z})}{\sum_{z=i}^{Z} A_z} \tag{6-9}$$

where $DT^{s}_{t,z}$ is the discomfort time according to ASHRAE55 at time-step t in zone z during season s.

6.3.1.1 Selection of Optimization Algorithm

Optimization problems can be mainly categorized as single objective or multi-objective optimization problems, where the former have only one objective function and the latter have more than one objective function. These objective functions are usually in conflict with each other in real-world engineering optimization problems so that the improvement of one of them leads to worsening the others. Therefore, multi-objective optimization offers the near-optimal set of solutions, which are called Pareto points or Pareto front, rather than a single near-optimal solution. In this set, there is not any answer that dominants the others (Deb, 2005).

Different analytics and heuristic optimization methods have been used by researchers in order to solve the optimization problems related to energy management objectives, such as improving energy efficiency, reducing energy cost, and increasing the occupants' comfort. The approximations (or heuristics) algorithms are used when finding exact optimal solutions is not applicable. Although heuristic optimization algorithms find approximate feasible solutions within a reasonable time frame, there is no guarantee of optimality. To overcome this shortcoming, meta-heuristic methods were developed that employ heuristics techniques with guidance through the search space to obtain near-optimal solutions (Mellouk et al., 2015).

Among different meta-heuristic optimization algorithms, the Genetic Algorithm (GA) followed by Particle Swarm Optimization (PSO) are the most used ones in the energy management field (Shaikh et al., 2016). This is due to the capability of GA in solving complex multi-objective optimization problems while maintaining the simplicity of its computational steps. GAs mimic the process of natural selection in order to find proper solutions to optimization problems based on the ideas of the evolutionary theory (Holland, 1975). PSO algorithm is an evolutionary computation technique, which is motivated by the behavior of bird flocks. Similar to GA, the PSO algorithm generates a population of random solutions called particles. However, unlike GA, each particle is also associated with a randomized velocity. Thus, particles fly around a multi-dimensional search space to find out optimal solutions (Shi, 2001; Sun et al., 2004). Based on the literature, while both PSO and GA obtain high-quality solutions, the number of computational steps for GA is lower than that of PSO, which is due to the communication between the particles after each generation (Panda and Padhy, 2008). The difference in computational effort between PSO and the GA is problem-dependent. PSO, in general, outperforms GA for unconstrained nonlinear problems with continuous design variables. However, when applied to highly nonlinear, constrained optimization problems, that are typical for complex energy management problems, GA is more efficient and requires less computational time (Hassan et al., 2005). Therefore, choosing an optimization algorithm with less computational steps, such as GA, would result in producing near-optimal solutions while reducing the complexity of the problem.

Among various multi-objective evolutionary algorithms (MOEAs), the Non-Dominated Sorting Genetic Algorithm (NSGA) was one of the first methods to create Pareto-optimal solutions (Srinivas and Deb, 1994). However, in order to alleviate some of the problems associated with NSGA, a better and faster algorithm, called NSGA-II, was introduced a few years later. Computational complexity, lack of elitism, and the need for sharing parameters were some of those problems (Deb et al., 2002). The new algorithm performs better and faster to find the non-dominated solutions by providing a better distribution of the population. According to (Wang, 2016) the NSGA-II is a mature multi-objective optimization algorithm at present. The main advantages of NSGA-II includes the flexibility to be applied to a wide range of optimization problems of significant complexity (McCall, 2005; Deb et al., 2002), the simplicity of its computational steps, especially when it is integrated with simulation models, and its ability to effectively solve multi-objective optimization problems. Therefore, NSGA-II is selected as the optimization engine in this research.

6.3.2 Simulation Model

In order to evaluate the energy performance of an open-plan office in terms of energy consumption and its occupants' discomfort time, a simulation model of a shared space is created as a basis to develop the proposed integrated model. Given outside weather conditions, building characteristics (e.g., building location, geometry, envelope, etc.), building systems' characteristics (e.g., the type of HVAC and lighting systems and their specifications), and internal loads (e.g., occupancy loads), the simulation model calculates the building energy consumption as well as the occupants' thermal condition mimicking the actual building energy performance. In this study, the simulation model is developed to investigate the effect of the application of multiple local control strategies on the energy performance of the building considering the dynamic occupancy information.

6.3.3 Integrated Simulation-based Optimization Model

Figure 6-2 depicts the integration procedure of the simulation and the optimization models. The optimization algorithm starts with creating the initial population of size H in the first generation. The optimization decision variables, as discussed in Section 6.3.1, are randomly varied within their pre-defined ranges to create the members of the population. Each member of the population, which contains a set of decision variables, is fed into the simulation model as part of its inputs. Based on the values of the decision variables in each simulation run, the objective functions are calculated. After calculating the fitness values of all members of the population, the selection, crossover, and mutation operations are performed on the entire population. This procedure is repeated for all members of the population in all generations until the convergence criterion (i.e., the specified number of generations, *M*) is met. After that, the optimization terminates and the optimal solutions are obtained as a Pareto front.

The output of the simulation-based optimization model is a file containing the information pertinent to the local control strategies. This file provides the following information: (1) the type of the control action, such as turning on/off the lights; and (2) the optimal settings for each zone of the space. By capturing the locations of the occupants at the zone level, the HVAC and lighting systems in the corresponding zone are activated and adjusted using the optimal results obtained from the integrated framework of the simulation-based optimization considering a trade-off between occupants' satisfaction and the energy consumption. It is important to mention that defining detailed control strategies regarding the building's energy-consuming systems are out of the scope of this research. In this research, control strategies are limited to changing the space temperature and turning on/off the lighting system using the thermostat setting and light switches, respectively.

6.3.4 HVAC and Lighting Operation Local Controls

- **HVAC System Local Control**

The operation of the HVAC system to reach the desired temperature is not an instantaneous process like other energy-consuming systems (e.g., light). Therefore, having information regarding the next state of a space usage plays an important role in the operation of the HVAC system. To this end, the dynamic space usage patterns are required. Although the techniques proposed in this area of research are well-established, only few studies considered the optimal operation of HVAC system based on dynamic occupancy information in shared open-plan offices (Goyal et al., 2013; Oldewurtel et al., 2013; Dobbsand Hencey, 2014a-b; Dong and Lam, 2014; Capozzoli et al., 2017).

In office buildings, the HVAC system usually runs based on pre-determined set-point and set-back temperatures during the occupied and unoccupied hours, respectively. This corresponds to control the temperature for defined fixed schedules (e.g., 8 am - 8 pm). Occupancy schedules, however, are highly dependent on occupants' work habits and they may deviate from one zone to another in the same shared office. Running the HVAC system based on a predefined schedule could result in higher energy consumption as well as occupants' dissatisfaction.

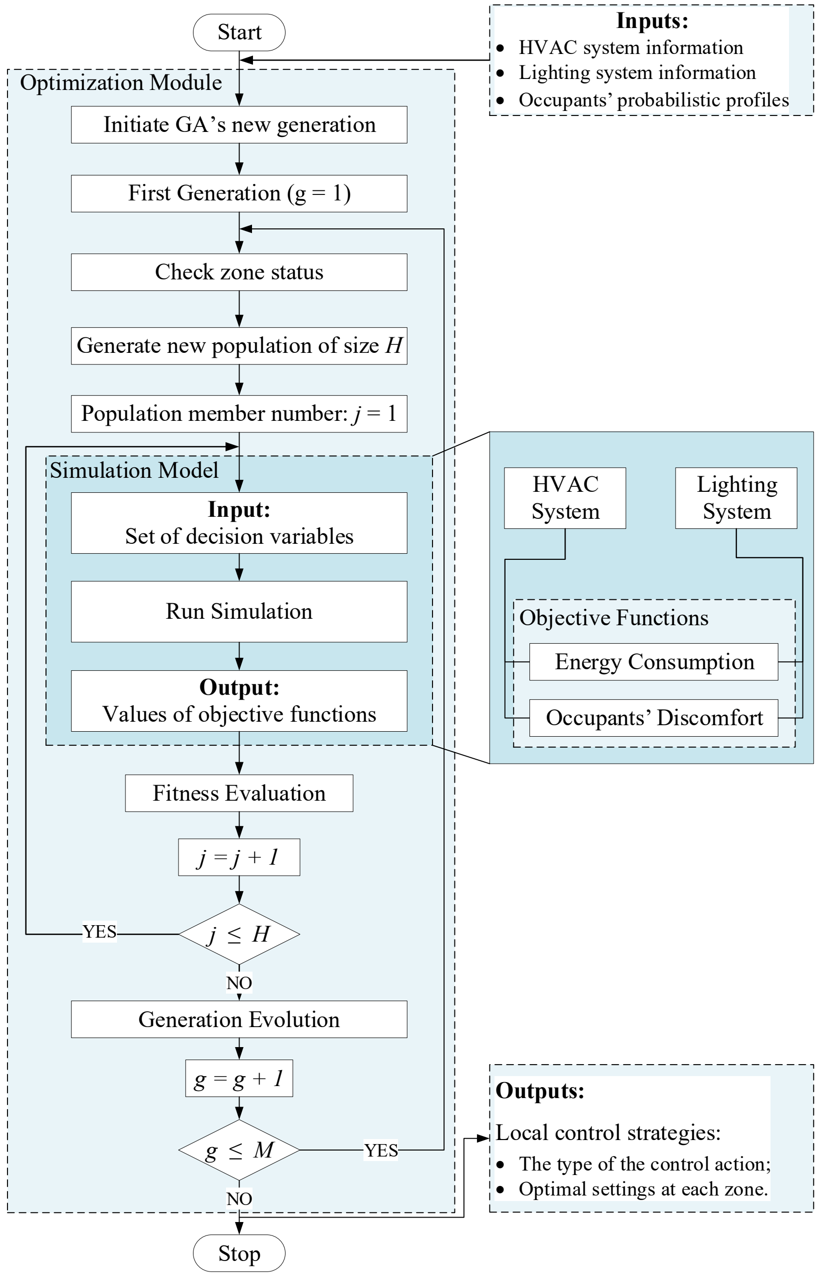


Figure 6-2 Integration of optimization and simulation models

For instance, if occupants leave their office earlier than the time in the fixed schedule, conditioning an empty space causes unnecessary energy consumption. On the other hand, working later than the pre-determined time in the fixed schedule in spaces conditioned based on the set-back temperature will have a negative impact on the occupants' productivity. Having prior information pertinent to occupancy changes at the zone level can be used as input for the optimal local control of the HVAC system (schedule-based HVAC local control). Figure 6-3 depicts the pseudocode showing how the HVAC local control algorithm adjusts the cooling set-point temperature during the cooling season using dynamic occupancy information. The same control procedure is used to set the heating set-point temperature with different sets of temperatures.

Set TOT_c= the total number of occupancy time-steps during the cooling period, m = time-step resolution to control the HVAC system, HS_c= HVAC cooling schedule, OCT_{sp}^{z} = cooling set-point temperature outside the comfort range, ICT_{sp}^{z} = cooling set-point temperature inside the comfort range, CT_{sb}^{z} = cooling setback temperature, $T_{c,t}^{occ,z}$ = occupied zone cooling temperature, $T_{c,t}^{unocc,z}$ = unoccupied zone cooling temperature, $CE_{t,z}^{S}$ = cooling energy at time-step *t* at zone *z* during season *s*.

```
During cooling season
  For each z in Z
    determine Status_{z, m} and OccR_{z, m}
    For each m in TOT_c
      if Status_{z, m} = 1:
        HS_c = 1
        if PMV index is outside the comfort range:
          T_{c,t}^{occ,z} = OCT_{sp}^z
        else
          T_{c,t}^{occ,z} = ICT_{sp}^z
          calculate cooling energy consumption (CE_{t,z}^S) due to occupancy internal gains using OccR_{z, m}
      else
        HS_c = 0.5
        T_{c,t}^{unocc,z} = CT_{sb}^z
    end
  end
```

Figure 6-3 Pseudocode for HVAC local control during the cooling season

Furthermore, two different HVAC schedules are considered in this study called simple and detailed schedules to apply optimal local control strategies. In the simple schedule, the HVAC system starts one hour before occupants' first arrival to each zone to bring the zone to the desired temperature. The zone temperature will remain at set-point temperature throughout the day and the HVAC control system will set the zone temperature to the set-back temperature one hour after the last departure of the occupants at that zone. The detailed HVAC schedule, however, follows the zone occupancy schedule. In this case, like the simple schedule, the HVAC system starts one hour before occupants' first arrival to each zone and remains on for the first hour. Considering the lag time of the HVAC system to adjust the zone temperature, a one-hour time-step is used to check the occupancy status and adjust the zone temperature accordingly. As such, if the zone is unoccupied for at least one hour, the set-back temperature is used. Otherwise, the set-point temperature is applied even when the zone is unoccupied. This procedure is repeated for the following hours until the end time of occupancy at each zone.

- **Lighting System Local Control**

Lighting system control is performed to maintain an optimum set-point illuminance level in occupied zones based on the dynamic occupancy information. Using schedule-based lighting system control and turning off the lights during unoccupied periods lead to a reduction in the internal heat gains and eventually to a decrease in total building energy consumption.

According to the dynamic occupancy information imported to the simulation model, the zone status can be determined at each time-step. Knowing the zone status, the lighting system control algorithm, as demonstrated in Figure 6-4, sets the lighting operation schedule based on the dynamic occupancy schedule and adjusts the lighting power according to the illuminance level of the occupied zone selected by the optimization algorithm.

Set LS = lighting operation schedule, i= occupancy time-step resolution, *TOT*= the total number of occupancy time-steps in a year.

During all seasons
  For each z in Z
    read *Status* $_{z,\,i}$ data series
    For each i in *TOT*
      if *Status* $_{z,\,i}$ = 1:
        Select randomly $I^{s}_{t,z}$ from decision variables table
        turn the lights on
        calculate the lighting power using Equation (6-8)
        $LP^{s}_{t\,,z}= LS \times LP^{s}_{t\,,z}$
      else
        turn the lights off
    end
  end

Figure 6-4 Pseudocode for lighting local control

6.3.5 Building Energy Management System (BEMS)

In order to integrate the dynamic occupancy-centered operation schedules of building energy-consuming systems with the simulation model, the EnergyPlus Energy Management System (EMS) is used. The EMS module is a high-level, generalized, supervisory control for building systems that utilizes EnergyPlus Runtime Language (Erl), which is a simplified programming language, to define the EMS control and modeling programs (DOE, 2015).

The EMS employs a set of *sensors* to retrieve information regarding the building's internal and external conditions. The collected data defines Erl variables to be used in control programs. Different changes are then performed on building systems through EMS *actuators*. Co-ordination between EMS *sensors* and *actuators* is performed through EnergyPlus simulation (DOE, 2015). Using EMS helps to emulate, inside EnergyPlus, the same control actions implement through BEMS in real buildings (Sehar et al., 2016; Tahmasebi and Mahdavi, 2018). Figure 6-5 illustrates the optimal local control algorithm used in this study. Although most of the traditional BEMSs control building systems using 15-minutes time-step intervals, with the growing advances in Information and Communication Technology (ICT) and the application of advanced data analytics to near real-time data gathered by Internet of Things (IoT) sensors, finer resolution levels are used by new BEMSs to monitor and control building systems (Sehar et al., 2017; Salimi and Hammad, 2018). This shows the importance of considering high-resolution levels for the application of control strategies. On the other hand, different resolution levels are required for controlling different building systems based on the discussion in Section 6.3.4. For instance, a higher level of resolution (i.e., time-step i in Figure 6-5) is needed to apply lighting control strategies, which improves the comfort level. However, considering the required lag time for the HVAC system to adjust the indoor temperature to a specified target set-point/set-back, a lower level of resolution (i.e., m in Figure 6-5) is needed to provide the required thermal comfort. As a result, different time-steps are defined to develop the dynamic occupancy profiles for the control of various building systems as explained in Section 6.2. As illustrated in Figure 6-5, time-step i is used to derive the zone occupancy patterns and accordingly adjust different schedules within the simulation model. This includes schedules regarding the zone occupancy, equipment, and lighting systems. If the zone is occupied, equipment and lights are turned on and the lighting power is calculated. Since i is smaller than m, the control module with time-step i is run for m/i times. Then, the HVAC system is controled using the derived dynamic occupancy profiles according to the procedure shown in Figure 6-1. In this study, i is considered as one minute. Two resolution levels of 30 and 60 minutes are used as m to control the HVAC system.

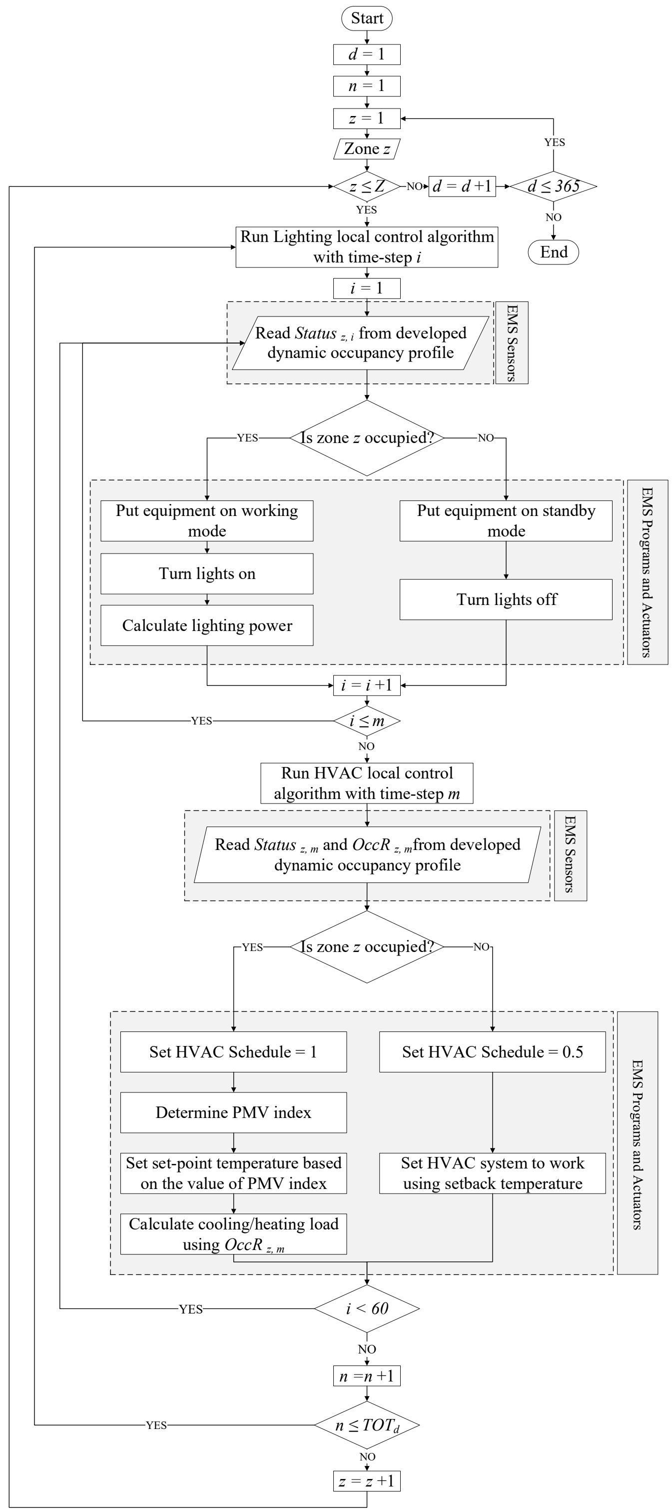


Figure 6-5 Optimal control flowchart for the operation of building systems

To run the optimal local control algorithm, the EMS *sensors* collect information pertinent to the occupancy status of each zone of the space (i.e., occupied or not in use) and the number of present occupants in the case of the occupied zone at each time-step. Based on the retrieved data, necessary control actions defined in EMS *programs* are performed through the EMS *actuators*. For instance, to apply HVAC system local control, once the EMS *sensors* retrieved the zone occupancy information, the optimization algorithm determines the optimal set-point temperature of the zone according to the values of the PMV index. Then, EMS actuators change the zone temperature to reach the optimal setting. At predefined calling points specified by EMS *ProgramCallingManager*, EMS *programs* are run during EnergyPlus simulation (DOE, 2015).

6.4 Case Study

6.4.1 Simulation Model and Input Data

In order to evaluate the energy performance of an open-plan office, a simulation model of a shared space is described in this section as a basis to develop the proposed integrated model. The simulation model is created in EnergyPlus version 8.6 (EnergyPlus, 2015). The model's layout is defined based on the plan of a real open-plan office. EnergyPlus, a powerful dynamic building energy simulation tool, offers accurate evaluation of the energy performance of buildings by predicting the dynamic behavior of building systems under ever-changing internal and external conditions (Sehar et al., 2017). Providing the building characteristics, such as building location and geometry, and its energy-consuming systems specifications, internal loads (e.g., occupancy, office equipment, and lighting) along with the external input (i.e., weather data), the simulation model calculates the space energy consumption, discomfort hours, and other required parameters at specified time-steps. Table 6-1 summarizes the input data of the simulation model.

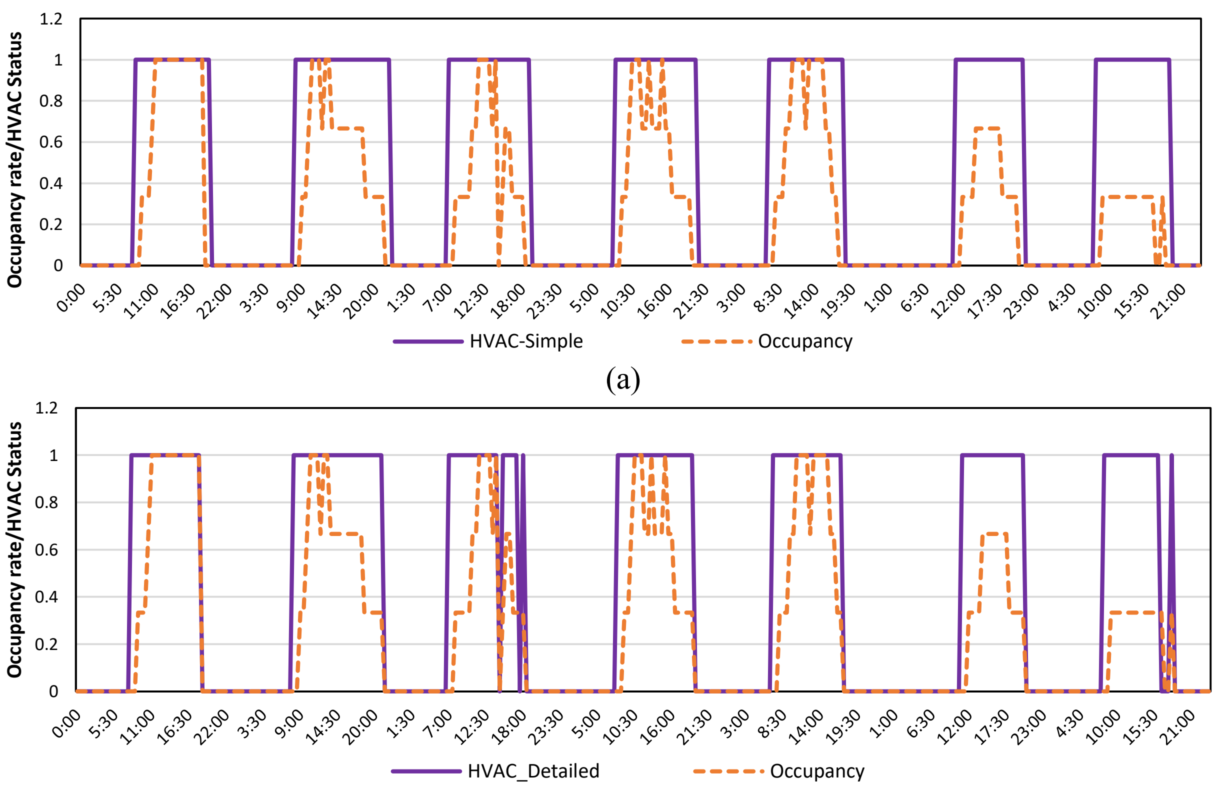
Table 6-1 Input data of the open-plan office simulation model

Simulation Model Inputs		Value
Space floor area (m^2)		35.1
Wall height (m)		3.5
Number of occupants at each zone		3
Office equipment power density (W/m^2) (Raji et al., 2017)		11.77
COP		5.5
Coefficient of efficiency		0.89
Lighting power density (W/m^2/100 lux)	Zone 1	1.38
	Zone 2	3.10

- Weather data: The weather data from Montréal–Pierre Elliott Trudeau International Airport available in (EnergyPlus, 2015) are used in this study.

- HVAC load: The normal heating and cooling set-points are 22 °C and 24 °C, respectively. The set-back temperatures are set to 18 °C and 28 °C during winter and summer seasons, respectively, for unoccupied periods for all zones according to the ASHRAE 90.1-2007 recommendations (ASHRAE, 2007). The simulated office has two zones.
- Occupancy, lighting, and equipment loads: There are three occupants located at each zone of the open-plan office. The occupancy profiles representing the working schedules of each occupant are derived from real occupancy data collected from the shared office for one year. The lighting and equipment schedules follow the occupancy profiles in this study. According to (Raji, Tenpierik, & van den Dobbelsteen, 2017), the office equipment power density is set to 11.77 (W/m^2). Since the number of luminaires varies between zones, each zone has its own lighting power density as shown in Table 6-2.

The occupancy and HVAC schedules are shown in Figure 6-6 for a one-week period.



(a)

(b)

Figure 6-6 Occupancy and HVAC system schedules during one summer week (a) Simple; (b) Detailed

Since the operation of building systems are highly dependent to the presence of occupants, the real occupancy data of the simulated open-plan office are collected over the course of one year (from April, 1st 2017 to March, 31st 2018) using Bluetooth Low Energy (BLE)-based monitoring system with a very high temporal resolution level (i.e., each second). High-resolution occupancy data are required for some specific applications, such as security and emergency situations. However, the high granularity of one second is not required for building energy management. Therefore, the real occupancy data are generated with the time intervals of one minute. After processing the raw collected data to have the occupants-specific dynamic profiles, the occupancy information at zone level with required resolution levels are imported to the simulation model in order to emulate the occupants' dynamic space utilization patterns. The integration of the dynamic occupancy-centered operation schedules of building systems derived from the processed raw data with the simulation model is performed using EnergyPlus EMS. As mentioned in Section 6.3.5, different resolution levels are considered for the simulation model's input data. Moreover, EnergyPlus EMS is integrated with an optimization algorithm to design different optimal local control strategies using jEPlus+EA (Zhang, 2009; Zhang, 2012), a third-party optimization tool developed for EnergyPlus. In order to apply the local control strategies of building systems based on dynamic occupancy information, the zoning is used to assign different dynamic occupancy-centered schedules for the operation of building systems at each zone. These zones are created using virtual partitions to separate open spaces without having physical boundaries that could affect the energy consumption of the space. During zone occupancy, set-point temperatures are adjusted according to Table 6-2 for cooling and heating set-points. These ranges are selected based on the average occupants' preferences in each zone. In case of turning the lights on, the illuminance level of the occupied zone is set to be between 300 to 500 lux recommended by Illuminating Engineering Society of North America (IESNA) for office buildings (DiLaura et al., 2011).

Table 6-2 Optimization variables

Variable	Min	Max	Increment
Cooling set points outside the comfort range (OCT_{sp}^{z})	26 ℃	27 ℃	0.5 ℃
Cooling set points inside the comfort range (ICT_{sp}^{z})	26 ℃	28 ℃	0.5 ℃
Heating set points outside the comfort range (OHT_{sp}^{z})	24 ℃	26 ℃	0.5 ℃
Heating set points inside the comfort range (IHT_{sp}^{z})	22 ℃	24 ℃	0.5 ℃
Illuminance level at each zone ($I_{t,z}^{s}$)	300 lux	500 lux	50 lux

6.4.2 Building Performance Metrics

To evaluate the performance of the simulation model under the application of different optimal local control strategies, the annual building energy consumption along with the time outside ASHRAE 55 comfort regions are considered as building performance indicators. The building performance indicators without the application of control strategies are used as the baseline and are denoted by P_0^{unc}. The building performance metrics are calculated as follows:

$$\Delta P = \frac{P^c - P_0^{unc}}{P_0^{unc}} \tag{6-10}$$

where P^c is the building performance indicators after the application of optimal local control strategies.

6.4.3 Results

All simulations are run using the 1-minute resolution level for a whole year. However, according to the discussion in Section 6.3.4, the HVAC system control is performed at 30- and 60-minute time-steps. Six different cases are investigated in this study among which four cases incorporate local control strategies using the optimization algorithm, and the other two cases are simulation models without the application of local control strategies nor optimization. Table 6-3 shows the specifications of the six cases.

Table 6-3 Specifications of investigated cases

Case No.	1	2	3	4	5	6
Occupancy schedule	Standard schedules	Real schedules	Real schedules	Real schedules	Real schedules	Real schedules
Occupancy resolution level	60 min	1 min	1 min	1 min	1 min	1 min
HVAC schedule	Standard schedules	Simple	Simple	Simple	Detailed	Detailed
Temperature	Fixed	Fixed	Variable	Variable	Variable	Variable
HVAC control resolution level	-	-	30 min	60 min	30 min	60 min
Lighting control resolution level	-	-	1 min	1 min	1 min	1 min
Optimization	-	-	✓	✓	✓	✓

In this table, standard schedules are standard office schedules with the maximum occupancy between 8:00 am and 5:00 pm and an hour of reduced occupancy for lunch at noon. Equipment and lighting are set to be on during occupancy periods in the first case. Occupancy, equipment and lighting schedules in the remaining cases, however, follow the real occupancy patterns derived from the real collected data. The heating and cooling set-point temperatures are fixed at 22 °C and 24 °C; respectively, for the first two cases. Figure 6-7 illustrates the sets of Pareto optimal solutions for the four optimization cases. In all these cases, optimization is run for 100 generations with a population size of 20. For the two resolution levels, cases with the simple HVAC schedule outperform the ones with the detailed schedule from the comfort point of view. This observation indicates that simpler schedules for operating the HVAC system result in fewer discomfort hours. That is mainly due to the lag time required for the HVAC system to adjust the indoor temperature. Frequent changes in the zone temperature result in more dissatisfied occupants, which subsequently leads to a decrease in their productivity. For instance, although solutions A and B have the same energy consumption, there is almost 11% increase in the discomfort hours as shown in Figure 6-7. On the other hand, cases with the detailed HVAC schedule generate more energy conservative solutions as the Pareto fronts of these cases are more towards the lower right side of the graph. Furthermore, the cases with lower resolution level for the application of local control strategies (i.e., 60 min) generate slightly better optimal solutions than those of obtained from the cases with more frequent application of local control (i.e., 30 min), such as solutions C and D in Figure 6-7. As a result, the simple HVAC schedule with the 60-minute resolution level (Case 4) found to be the best option for the application of local control strategies.

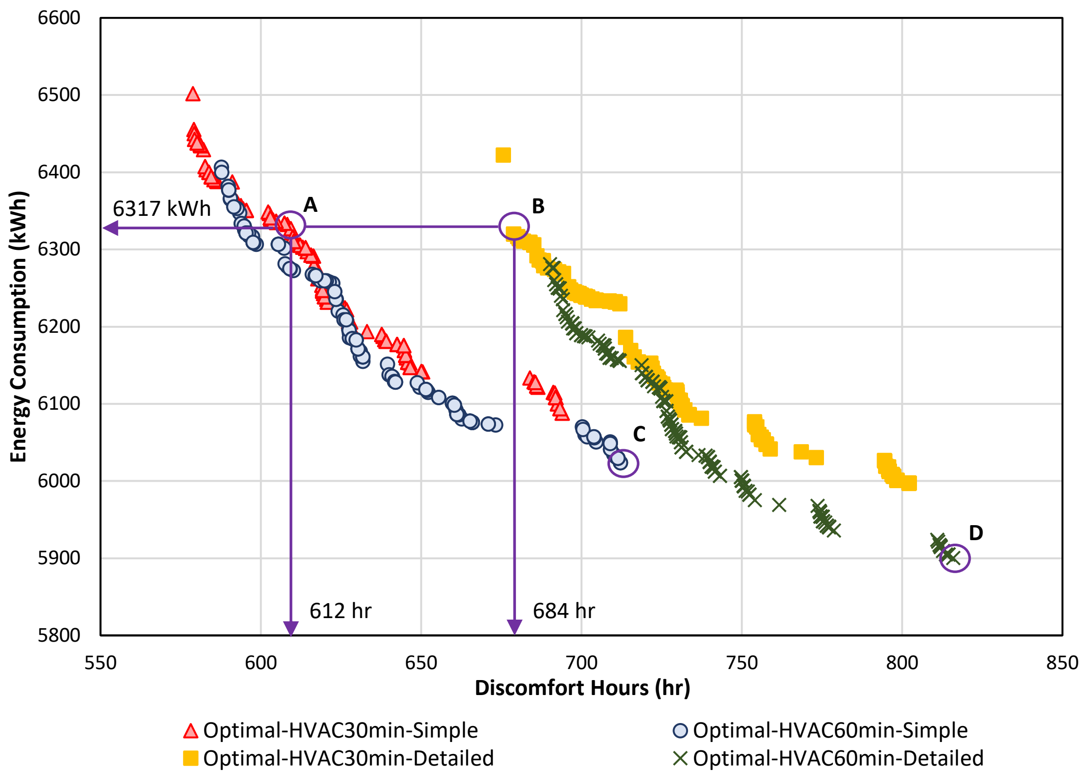


Figure 6-7 Sets of Pareto solutions for four optimization cases

The Pareto solutions Case 4 along with all other solutions generated by the optimization algorithm are shown in Figure 6-8. Almost 8% of the results of the optimization are optimal solutions that make the Pareto front as demonstrated in this figure. According to Figure 6-8, three distinct areas are distinguished as shown by square, ellipse, and circle areas. The optimal solutions within the square area represent solutions with lower discomfort hours and higher energy consumption. The optimal solutions enclosed within the circle area, however, show the opposite trend (i.e., higher discomfort hours and lower energy consumption). Choosing any of these solutions means preferring one objective function for achieving better results at the expense of the other objective function. Therefore, the most desirable solutions that correspond to a trade-off between the two objective functions are the solutions inside the ellipse area. For Case 4, the most desirable solutions account for almost 51% of solutions (76 out of 150 optimal solutions). The average discomfort hours and energy consumption associated with this area are about 645 hours and 6138 kWh, respectively.

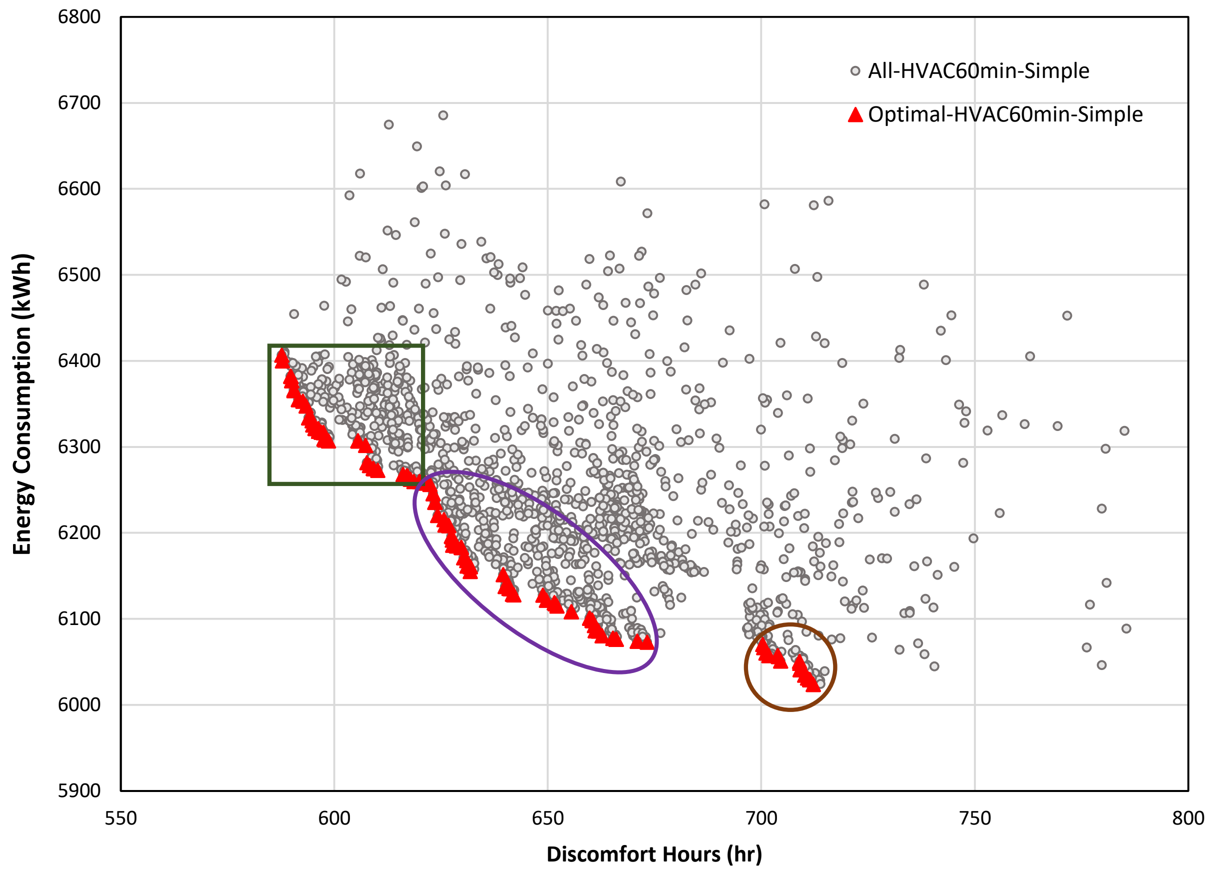


Figure 6-8 Optimization results for Case 4

The comparison of the average discomfort hours and energy consumption of the most desirable optimal solutions in the four optimal cases are shown in Table 6-4. The information regarding these solutions is also included in this table. Case 6 generates the best solutions in terms of building energy consumption. However, the highest number of discomfort hours is reported in this case. A balance between the two objective functions is observed in Case 4. This conclusion is aligned with the observations related to the whole Pareto optimal solutions (Figure 6-7).

Table 6-4 Average of most desirable optimal solutions for four optimal cases

Case No.	**No. of optimal solutions (Percentage of total solutions %)**	**No. of most desired solution (Percentage of optimal solutions %)**	**Total energy consumption (kWh)**	**Discomfort hours (hr)**
3	191 (10)	110 (58)	6227	628
4	150 (8)	76 (51)	6138	645
5	233 (12)	59 (25)	6116	728
6	243 (13)	104 (43)	6076	729

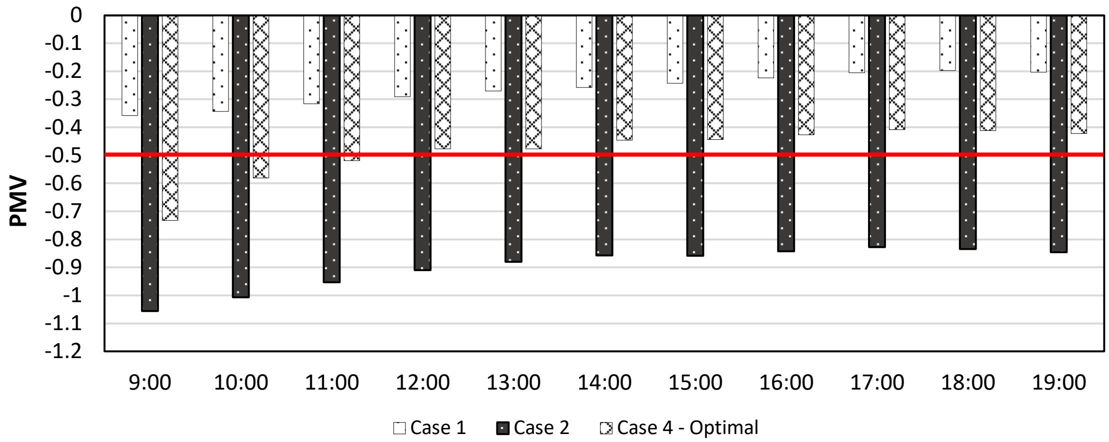
After selecting Case 4 as the best case for the application of local control strategies, the results of comparing this case with the first two cases are shown in Table 6-5. The comparison demonstrates that although using standard schedules leads to fewer discomfort hours, there is 15% more energy consumption compared to optimal solutions of Case 4. On the other hand, following real occupancy schedules without changing the set-points results in a considerable increase in the number of total discomfort hours (more than 148%) compared to standard schedules. Using optimization helps to find a balance between these two extreme cases. Although the improvement in the office energy consumption (i.e., 2%) is not as much as that of the discomfort hours, these hours could be cut by more than 50% when comparing Cases 2 and 4.

Table 6-5 Comparison among three cases

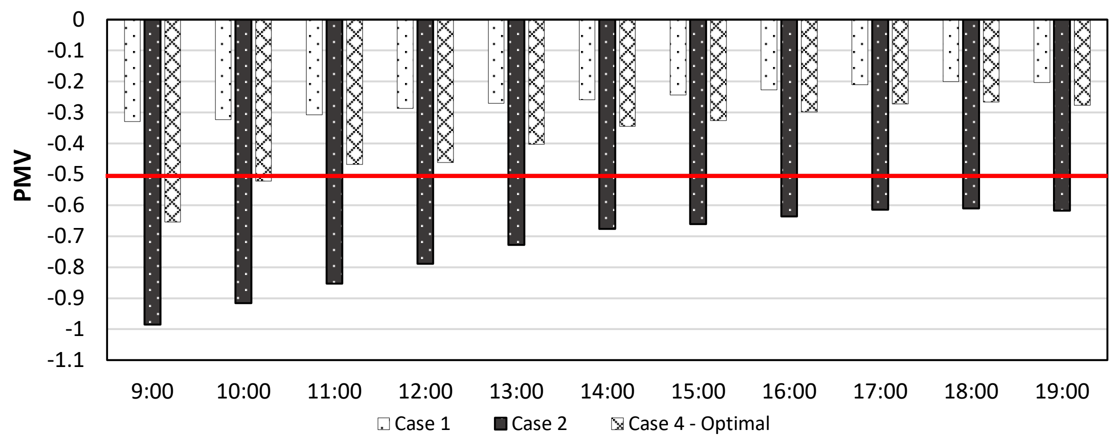
Case No.	Total energy consumption (kWh)	Performance metric (%)	Discomfort hours (hr)	Performance metric (%)
1	7065	-15	395	+39
2	6266	-2	982	-52
4	6138	-	645	-

The comparison between the PMV index of the three Cases 1, 2, and 4 in Zones 1 and 2 is demonstrated in Figure 6-9 for a winter day in January. Moreover, Figure 6-10 illustrates the changes in the PMV index during a summer day in June. The outcomes of one optimal solution, as a representative of the solutions enclosed within the ellipse area in Case 4 (as shown in Figure 6-8), are used for the comparison purpose.

As shown in Figure 6-9, Case 1 demonstrates better performance compared to the other two cases from the comfort point of view during winter. However, comparing Cases 2 and 4 shows that the PMV index remains greater than -0.5 for the optimal case compared to that of Case 2 during most of the occupied hours except the start time of occupancy. Although the HVAC system is set to bring the zone temperature to the set-point one hour before the start time of occupancy, this may not be always possible. This problem can be solved by considering earlier wake-up time for the HVAC system to provide enough time to adjust the zone temperature.



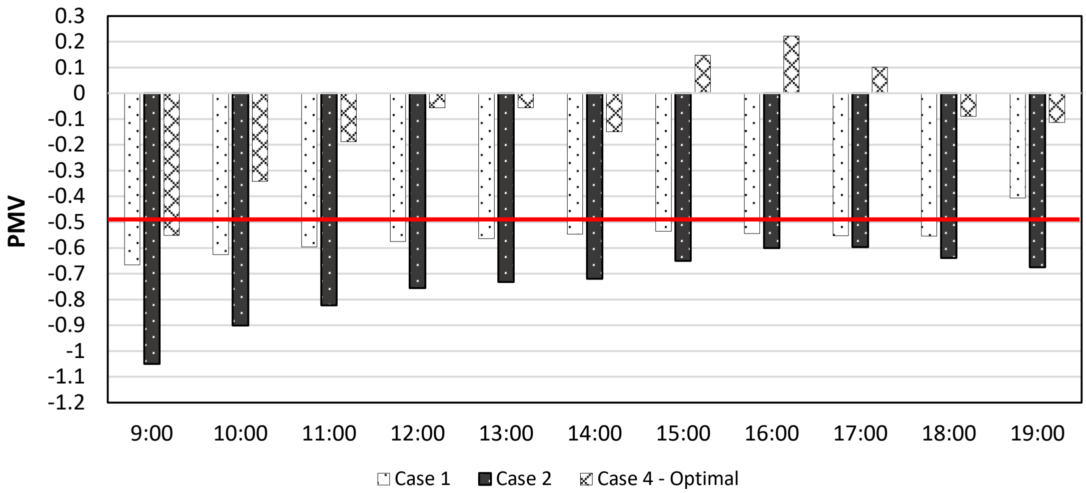
(a)



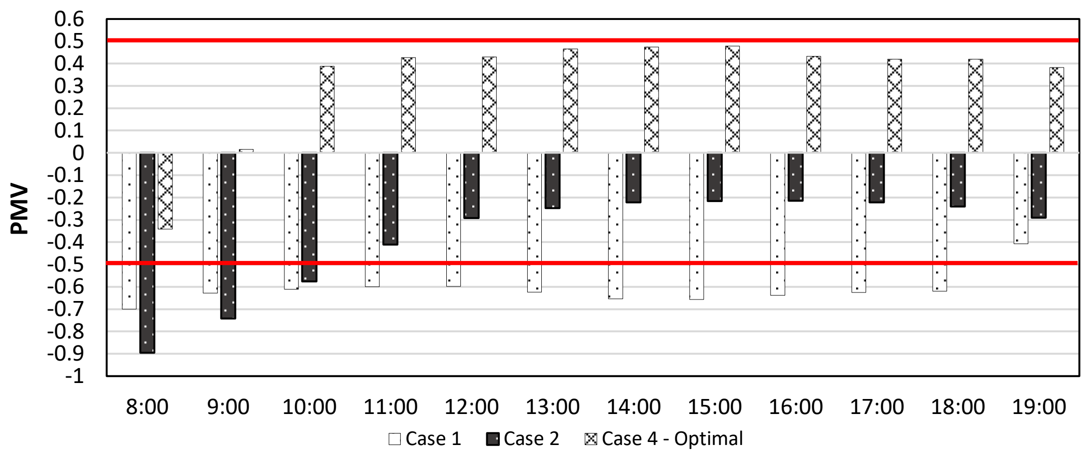
(b)

Figure 6-9 Changes in PMV index during a winter day (a) Zone 1; (b) Zone 2

As opposed to the heating season, analyzing the obtained results from Case 1 demonstrates that the most discomfort hours are happening during the cooling season. Therefore, the optimal case outperforms the other two non-optimal cases during most of the summer days. In Case 4, shifting the cooling temperature set-point by 3 ℃ (i.e., from 24 to 27 ℃) results in increasing the PMV index and improving the comfort condition. In addition, the PMV index in Case 4 becomes positive in many hours during the summer day showing the higher potential for saving energy while maintaining occupants' comfort during the cooling season compared to the wintertime.



(a)



(b)

Figure 6-10 Changes in PMV index during a summer day (a) Zone 1; (b) Zone 2

6.5 Summary and Conclusions

Occupancy-related information is recognized as influential factors affecting the energy operation in buildings. In order to enhance the energy management of buildings, occupancy-centered control strategies are required. Therefore, the need for localized and customizable comfort controls is increasing in open-plan office buildings to improve the occupants' satisfaction, and consequently their productivity. The main contribution of this study is optimizing energy consumption and occupants' comfort in open-plan offices using local control based on occupancy dynamic data. In this research, full-year dynamic occupancy profiles with high resolution of 1-minute are developed using real occupancy data capturing the stochastic behavior of occupants. Having the dynamic occupancy profiles along with the indoor environmental conditions, this study contributes to the exploration of optimal local control strategies based on the integration of a simulation model with a multi-objective optimization process. It is concluded that the simple HVAC schedule with the 60-minute resolution level is the best option for the application of local control strategies from the comfort point of view for the case study of this research. The application of HVAC and lighting local control strategies results in improving the thermal condition by 50% along with 2% savings in energy consumption. More energy conservative solutions are, however, generated upon the usage of detailed HVAC schedules for both resolution levels.

Using the proposed method in open-plan offices results in: (1) developing high-resolution dynamic occupancy profiles to represent temporal variations of occupancy patterns in open-plan offices; (2) developing local control algorithms for building energy-consuming systems; (3) finding the best settings to operate HVAC system including the type of HVAC schedule and the resolution level used to control this system; and (4) finding the trade-off between buildings' energy consumption and occupants' comfort levels by maintaining the PMV index within the comfortable range. The practical application of the proposed method aims at assisting decision-makers in evaluating optimized occupancy-centered building operations.

Since the air flows from the terminal units have different temperatures, the effect of air mixing should be considered. It is assumed that the air mixing of zones only happens in the boundaries. This assumption makes the proposed method more applicable in large open-plan offices, where air mixing effects do not considerably affect the zone temperatures. Thus, the effect of air mixing should be considered as future work. To do so, computational fluid dynamics (CFD) analysis of the proposed optimal local control strategies should be conducted to capture the air mixing between zones and the resultant variations in the zone temperature. In addition, the effect of using occupancy information generated by an occupancy prediction model should be investigated. To this aim, occupancy information generated by a probabilistic occupancy model can be fed to the integrated simulation-based optimization model for investigating the energy savings corresponding to predicted occupancy profiles. Analyzing the outcomes of this integration helps to evaluate the effect of occupancy prediction accuracy on the performance of the integrated model. Future work will also consider discomfort time from the lighting point of view including the effects of glare and positions of occupants with respect to windows.

CHAPTER 7 SUMMARY, CONTRIBUTIONS, AND FUTURE WORK

7.1 Summary of Research

This research covered a comprehensive review of the related literature, the current research gaps, the overview of the proposed framework, and detailed explanation of the proposed methods followed by the case studies to validate and evaluate the applicability of the proposed framework. In the literature review, different occupancy monitoring techniques, occupancy modeling approaches, and control systems for building energy-consuming systems were discussed. Moreover, a roadmap regarding the advances in different dimensions with respect to office buildings' energy management is presented. The proposed roadmap provides a high-level view of the directions for future research towards CBM. By integrating all the components in the roadmap, a vision of CBM can be seen where buildings' systems, their occupants, and all other stakeholders have intelligent support from systems encapsulating sensor data and control strategies.

In the proposed framework of this research, the methodology regarding the development of a new adaptive probabilistic occupancy model and a simulation-based multi-objective optimization framework was introduced. After performing the occupancy behavior analytics (data processing), important occupancy features, such as the number of present occupants, periods of absence and presence, and other occasional variations in the occupants' profiles are determined. Using the derived information, the proposed inhomogeneous Markov chain occupancy perdition model generates the probabilistic profiles of each specific occupant. The work state of each occupant, his/her location and the total number of present occupants can be derived from these profiles at each time-step.

Furthermore, after developing the energy simulation model of the office building, an integration framework of the simulation model with the optimization algorithm was proposed in order to improve the performance of the simulation model in evaluating the space energy usage. The optimization algorithm was designed to satisfy the two objective functions of minimizing the office building's energy consumption and occupants' discomfort. By capturing the locations of the occupants at the zone level, the HVAC and lighting systems in the corresponding zone are activated and adjusted using the optimal results obtained from the integrated framework considering a trade-off between occupants' satisfaction and the energy consumption.

7.2 Research Contributions and Conclusions

This research results in the following contributions:

(1) Developing a method for extracting detailed occupancy information with varying time-steps from collected RTLS occupancy data;

(2) Developing a new adaptive (self-learning) probabilistic occupancy prediction model based on the RTLS data to distinguish between different occupants within open-plan offices. With regard to the first two contributions the following conclusions can be drawn:

- The occupancy perdition model was able to accurately estimate occupancy patterns of the open-plan office at occupant and zone levels.
- The proposed prediction model is an adaptive model that evolves and improves itself over time.
- High accuracy of occupancy patterns prediction (86% and 68% on average for the purpose of the lighting and HVAC systems control, respectively) indicates the acceptable performance of the prediction model in capturing the temporal behavior of different occupants working in the same open-plan office.

(3) Improving the performance of the developed occupancy prediction model by applying sensitivity analyses. This contribution leads to the following conclusions:

- Months of data collection can be ranked using the proposed ranking procedure based on the data spread feature, the reliability, and the similarity between collected time-series.
- The application of the proposed sensitivity analyses determines the near-optimum length of the data collection period required at each zone of space along with the near-optimum training dataset.
- The performance evaluation also finds the most satisfying temporal resolution level for analyzing the occupancy data assuring acceptable accuracy in occupancy prediction.
- The required resolution for an acceptable accuracy level is dependent on the quality of the collected data.

- Obvious differences in the number of occupants and patterns in different zones indicate the importance of considering the occupancy data at the zone level instead of the room level.

(4) Developing local control algorithms for building energy-consuming systems.

(5) Developing a simulation-based multi-objective optimization model and assessing the effect of different intelligent and occupancy-centered local control strategies on building energy-consuming systems, which eventually leads to the application of different occupants' preferences with respect to building systems. The following conclusions are achieved from the last two contributions:
    - The proposed method improves the energy management of buildings by developing intelligent, optimal, and occupancy-centered local control strategies and evaluating the effect of them on building energy-consuming systems and the occupants' satisfaction.
    - The simple HVAC schedule with the 60-minute resolution level is the best option for the application of local control strategies from the comfort point of view.
    - The application of the simple HVAC schedule with the 60-minute resolution level and lighting local control strategies results in improving the thermal condition by 50% along with 2% savings in energy consumption.
    - More energy-saving solutions were, however, generated upon the usage of detailed HVAC schedules for both resolution levels of 30 and 60 minutes.

7.3 Limitations and Future Work

In spite of the above-mentioned contributions, there are limitations in this research that are needed to be addressed in the future. These limitations are as follows:

(1) Although the overall performance of the prediction model was satisfactory, it may not capture variations in occupancy patterns that may happen after the data collection period, especially in the case of open-plan offices with varying occupancy. This limitation could be solved by:
    - Having access to real-time occupancy detection and control.
    - Using different data collection periods and frequently updating the prediction model whenever a real-time occupancy detection and control happened to consider changes in the space utilization patterns. These considerations make the proposed prediction model more general for different types of open-plan offices.

- Using more advanced estimation approaches, such as machine learning techniques.

(2) In terms of performance analyses, a cost-benefit analysis should be done to investigate the balance between the cost of using different RTLSs, collecting data over a long period, and the gains of using them in the real world. In addition, the ranking procedure (proposed in Chapter 5) requires improvements to better fit its application for occupancy detection in open-plan offices. To do so:

   - Seasonal changes in occupancy patterns should be studied in detail by having different training datasets for different seasons.
   - Other methods, such as clustering, can be used for ranking the months in the training dataset.

(3) Since the air flows from the terminal units have different temperatures, the effect of air mixing should be considered. To address this issue, computational fluid dynamics (CFD) analysis of the proposed optimal local control strategies should be conducted to capture the air mixing between zones and the resultant variations in the zone temperature.

(4) In this study, occupants' comfort was studied by focusing on their thermal comfort. The discomfort time from the lighting point of view should be included to have a complete vision regarding occupants' discomfort time. It is recommended in the future to include the effects of glare and positions of occupants with respect to windows in order to account for discomfort time from the lighting perspective.

(5) There is a privacy issue when the occupants' identities are used to have detailed occupancy information. In the future, it is necessary to:

   - Anonymize the occupants' data through defining occupancy profiles per zone.
   - Clarify the importance of collecting this type of data for other purposes, such as emergency and safety.
   - Inform the monitored occupants about all the benefits coming from using the real-time monitoring system for a reasonable period.

(6) Future applications of emerging ICT in the building sector are towards real-time energy management. Thus, the ultimate goal of this research is to apply the near real-time occupancy responsive local control strategies on building energy-consuming systems in order to have responsive operational systems, which can learn and self-tune themselves for optimum and intelligent operation. In this matter:

- The results of the occupancy prediction model should be fed to the integrated simulation-based optimization model to perform near real-time energy management and investigate the energy savings corresponding to predicted occupancy profiles. Analyzing the outcomes of this integration helps to evaluate the effect of occupancy prediction accuracy on the performance of the integrated model.
- By feeding the occupancy pattern of space along with the local control strategies, derived from the occupancy prediction model and the simulation-based optimization module, respectively; the integrated model predicts the zone condition in near real-time and applies relevant local control strategies.

REFERENCES

Agarwal, Y., Balaji, B., Gupta, R., Lyles, J., Wei, M., & Weng, T. (2010). Occupancy-driven energy management for smart building automation. *In Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, (pp. 1-6).

Aghemo, C., Blaso, L., & Pellegrino, A. (2014). Building automation and control systems: A case study to evaluate the energy and environmental performances of a lighting control system in offices. *Automation in Construction, 43*, 10-22.

Ai, B., Fan, Z., & Gao, R. X. (2014). Occupancy estimation for smart buildings by an auto-regressive hidden Markov model. *In American Control Conference (ACC)* (pp. 2234-2239). Portland, Oregon, USA: IEEE.

Akanmu, A., Rasheed, S. H., & Qader, I. A. (2013). Spatial mapping approach to component tracking using RTLS system. *Architectural Engineering Conference 2013* (pp. 363-37). ASCE.

Al Bustani, H. (2014). *UAE sustainable skyscrapers: understanding Abu Dhabi's Al Bahar Towers*. (Abu Dhabi Investment Council) Retrieved from https://www.thenational.ae/business/uae-sustainable-skyscrapers-understanding-abu-dhabi-s-al-bahar-towers-1.248717

Al-Mumin, A., Khattab, O., & Sridhar, G. (2003). Occupants' behavior and activity patterns influencing the energy consumption in the Kuwaiti residences. *Energy and buildings, 35*(6), 549-559.

Amato, G., Carrara, F., Falchi, F., Gennaro, C., Meghini, C., & Vairo, C. (2017). Deep learning for decentralized parking lot occupancy detection. *Expert Systems with Applications, 72*, 327-334.

Analytics, G. (2015). *Seeing Quuppa's indoor location technology at MWC 2015*. Retrieved 4 15, 2017, from http://grizzlyanalytics.blogspot.ca/2015/03/seeing-quuppas-indoor-location.html

Annex 53. (2016). *Total Energy Use in Buildings: Analysis and Evaluation Methods.* AECOM Ltd.

Arora, A., Amayri, M., Badarla, V., Ploix, S., & Bandyopadhyay, S. (2015). Occupancy estimation using non intrusive sensors in energy efficient buildings. *Proceedings of BS2015: 14th Conference of International Building Performance Simulation Association* (pp. 1441-1448). Hyderabad, India: Building simulation.

ASHRAE. (2002). *ASHRAE Guideline 14-2002, measurement of energy and demand savings.* ISSN1049-894X. American Society of Heating Refrigeration and Air Conditioning Engineers.

ASHRAE. (2007). *Energy Standard for Buildings Except Low-Rise Residential Buildings.* ASHRAE, 90.1-2007.

ASHRAE/ANSI Standard 55-2010. (2010). *Thermal Environmental Conditions for Human Occupancy.* Atlanta, GA: American Society of Heating, Refrigerating, and Air-Conditioning Engineers.

Aswani, A., Master, N., Taneja, J., Culler, D., & Tomlin, C. (2012). Reducing transient and steady state electricity consumption in HVAC using learning-based model-predictive control. *Proceedings of the IEEE, 100*(1), 240-253. doi:10.1109/JPROC.2011.2161242

Attar, R., Hailemariam, E., Breslav, S., Khan, A., & Kurtenbach, G. (2011). Sensor-enabled Cubicles for Occupant-centric Capture of Building Performance Data. *ASHRAE Transactions, 117*(2), 441-449.

Augello, A., Ortolani, M., Re, G. L., & Gaglio, S. (2011). Sensor mining for user behavior profiling in intelligent environments. In V. Pallotta, A. Soro, & E. Vargiu, *Advances in Distributed Agent-Based Retrieval Tools* (pp. 143-158). Springer Berlin Heidelberg.

Azar, E., & Menassa, C. C. (2011). Agent-based modeling of occupants and their impact on energy use in commercial buildings. *Journal of Computing in Civil Engineering, 26*(4), 506-518.

Azar, E., & Menassa, C. C. (2012). A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings. *Energy and Buildings, 55*, 841-853.

Azzouzi, S., Cremer, M., Dettmar, U., Knie, T., & Kronberger, R. (2011). Improved AoA based localization of UHF RFID tags using spatial diversity. *RFID-Technologies and Applications (RFID-TA)*, (pp. 174-180).

Balaji, B., Xu, J., Nwokafor, A., Gupta, R., & Agarwal, Y. (2013). Sentinel: occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings. *In Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems.* Roma, Italy: ACM.

Benezeth, Y., Laurent, H., Emile, B., & Rosenberger, C. (2011). Towards a sensor for detecting human presence and characterizing activity. *Energy and Buildings, 43*(2), 305-314.

Bengea, S. C., Kelman, A. D., Borrelli, F., Taylor, R., & Narayanan, S. (2014). Implementation of model predictive control for an HVAC system in a mid-size commercial building. *HVAC&R Research, 20*(1), 121-135.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning, 2*(1), 1-127.

Bloomberg. (2015, September 23). *The Smartest Building in the World*. (PLP Architecture) Retrieved from https://www.bloomberg.com/features/2015-the-edge-the-worlds-greenest-building/

Bluetooth. (2016). *Bluetooth Core Specification v5.0.* Bluetooth SIG Working Groups.

Brackney, L. J., Florita, A. R., Swindler, A. C., Polese, L. G., & Brunemann, G. A. (2012). Design and performance of an image processing occupancy sensor. *In Proceedings: The Second International Conference on Building Energy and Environment 2012987 Topic 10* (pp. 987-994). Intelligent buildings and advanced control techniques.

Brager, G., Paliaga, G., & De Dear, R. (2004). Operable Windows, Personal Control and Occupant Comfort. *ASHRAE*, pp. 17-35.

Brooks, J., & Barooah, P. (2014). Energy-efficient control of under-actuated HVAC zones in buildings. *In American Control Conference (ACC)* (pp. 424-429). Portland, Oregon, USA: IEEE.

Brooks, J., Goyal, S., Subramany, R., Lin, Y., Middelkoop, T., Arpan, L., . . . Barooah, P. (2014). An experimental investigation of occupancy-based energy-efficient control of commercial building indoor climate. *In Decision and Control (CDC)* (pp. 5680-5685). Los Angeles, California, USA: IEEE 53rd Annual Conference.

Brooks, J., Kumar, S., Goyal, S., Subramany, R., & Barooah, P. (2015). Energy-efficient control of under-actuated HVAC zones in commercial buildings. *Energy and Buildings, 93*, 160-168.

buildingSMART. (2018). *Technical Vision*. Retrieved from buildingSMART International home of openBIM: https://www.buildingsmart.org/standards/technical-vision/

Caicedo, D., & Pandharipande, A. (2016). Daylight and occupancy adaptive lighting control system: An iterative optimization approach. *Lighting Research & Technology, 48*(6), 661-675.

Caicedo, D., Li, S., & Pandharipande, A. (2017). Smart lighting control with workspace and ceiling sensors. *Lighting Research and Technology, 49*, 446–460. doi:1477153516629531

Caicedo, D., Pandharipande, A., & Vissenberg, M. C. (2015). Smart modular lighting control system with dual-beam luminaires. *Lighting Research & Technology, 47*(4), 389-404.

CapotaLand. (2017). *The future of urban living*. (CapitaLand) Retrieved from https://www.capitaland.com/sg/en.html?id=11

Capozzoli, A., Piscitelli, M. S., Gorrino, A., Ballarini, I., & Corrado, V. (2017). Data analytics for occupancy pattern learning to reduce the energy consumption of HVAC systems in office buildings. *Sustainable cities and society, 35*, 191-208.

Chang, W. K., & Hong, T. (2013). Statistical analysis and modeling of occupancy patterns in open-plan offices using measured lighting-switch data. *In Building Simulation, 6*(1), 23-32.

Chen, J., & Ahn, C. (2014). Assessing occupants' energy load variation through existing wireless network infrastructure in commercial and educational buildings. *Energy and Buildings, 82*, 540-549.

Chen, Z., Xu, J., & Soh, Y. C. (2015). Modeling regular occupancy in commercial buildings using stochastic models. *Energy and Buildings, 103*, 216-223.

Cho, Y., Lim, S. O., & Yang, H. S. (2010). Collaborative occupancy reasoning in visual sensor network for scalable smart video surveillance. *IEEE transactions on Consumer Electronics, 56*(3), 1997-2003.

Christ, M., Kempa-Liehr, A. W., & Feindt, M. (2016). Distributed and parallel time series feature extraction for industrial big data applications. *arXiv preprint arXiv:1610.07717.*

Chung, T. M., & Burnett, J. (2001). On the prediction of lighting energy savings achieved by occupancy sensors. *Energy engineering, 98*(4), 6-23.

Çiftler, B. S., Dikmese, S., Güvenç, I., Akkaya, K., & Kadri, A. (2017). Occupancy counting with burst and intermittent signals in smart buildings. *IEEE Internet of Things Journal, 5*(2), 724-735.

Clevenger, C. M., & Haymaker, J. (2006). The impact of the building occupant on energy modeling simulations. *Joint International Conference on Computing and Decision Making in Civil and Building Engineering*, (pp. 1-10). Montreal, Canada.

Conte, G., De Marchi, M., Nacci, A. A., Rana, V., & Sciuto, D. (2014). BlueSentinel: a first approach using iBeacon for an energy efficient occupancy detection system. *In BuildSys@ SenSys* (pp. 11-19). Memphis, TN, USA: ACM.

Curran, K., Furey, E., Lunney, T., Santos, J., Woods, D., & McCaughey, A. (2011). An evaluation of indoor location determination technologies. *Journal of Location Based Services, 5*(2), 61-78.

D'Oca, S., & Hong, T. (2015). Occupancy schedules learning process through a data mining framework. *Energy and Buildings, 88*, 395-408.

Daum, D., & Morel, N. (2010). Assessing the total energy impact of manual and optimized blind control in combination with different lighting schedules in a building simulation environment. *Journal of Building Performance Simulation, 3*(1), 1-16.

Davis, J. A., & Nutter, D. W. (2010). Occupancy diversity factors for common university building types. *Energy and buildings, 42*(9), 1543-1551.

Day, J. K., & Gunderson, D. E. (2015). Understanding high performance buildings: The link between occupant knowledge of passive design systems, corresponding behaviors, occupant comfort and environmental satisfaction. *Building and Environment, 84*, 114-124.

de Bakker, C., Aries, M., Kort, H., & Rosemann, A. (2017). Occupancy-based lighting control in open-plan office spaces: A state-of-the-art review. *Building and Environment, 112*, 308-321.

Deb, K. (2005). Muti-Objective Optimization. In E. Burke, & G. Kendall, *Search Methodologies Introductory Tutorials in Optimization and Decision Support Techniques* (pp. 97-125). United States of America: Springer Science-i-Business Media, LLC.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions, 6*(2), 182-197.

Dedesko, S., Stephens, B., Gilbert, J. A., & Siegel, J. A. (2015). Methods to assess human occupancy and occupant activity in hospital patient rooms. *Building and Environment, 90*, 136-145.

Delaney, D. T., O'Hare, G. M., & Ruzzelli, A. G. (2009). Evaluation of energy-efficiency in lighting systems using sensor networks. *In Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings* (pp. 61-66). ACM.

Delgarm, N., Sajadi, B., Azarbad, K., & Delgarm, S. (2018). Sensitivity analysis of building energy performance: A simulation-based approach using OFAT and variance-based sensitivity analysis methods. *Journal of Building Engineering , 15*, 181-193.

Delgoshaei, P., Heidarinejad, M., Xu, K., Wentz, J. R., Delgoshaei, P., & Srebric, J. (2017). Impacts of building operational schedules and occupants on the lighting energy consumption patterns of an office space. *In Building Simulation, 10*(4), 447-458.

Deng, Z., & Chen, Q. (2019). Simulating the Impact of Occupant Behavior on Energy Use of HVAC Systems by Implementing a Behavioral Artificial Neural Network Model. *Energy and Buildings, 198*, 216-227.

Díaz, P., Olivares, T., Galindo, R., Ortiz, A., Royo, F., & Clemente, T. (2011). The ecosense project: an intelligent energy management system with a wireless sensor and actor network. *In Sustainability in Energy and Buildings* (pp. 237-245). Berlin Heidelberg: Springer.

Dikel, E. E., Newsham, G. R., Xue, H., & Valdés, J. J. (2017). Potential energy savings from high-resolution sensor controls for LED lighting. *Energy and Buildings, 156*, 43-53.

DiLaura, D. L., Houser, K. W., Mistrick, R. G., & Steffy, G. R. (2011). *The lighting handbook: Reference and application (10th ed.).* New York: Illuminating Engineering Society of North America (IESNA).

Dobbs, J. R., & Hencey, B. M. (2014a). Predictive HVAC control using a Markov occupancy model. *In American Control Conference (ACC)* (pp. 1057-1062). Portland, Oregon, USA: IEEE.

Dobbs, J. R., & Hencey, B. M. (2014b). Model predictive HVAC control with online occupancy model. *Energy and Buildings, 82*, 675-684.

Dodier, R. H., Henze, G. P., Tiller, D. K., & Guo, X. (2006). Building occupancy detection through sensor belief networks. *Energy and buildings,38(9),, 38*(9), 1033-1043.

DOE. (2015). *Application Guide for EMS Energy Management System–User Guide.* U.S. Department of Energy.

DOE. (2016). *DOE-2 Building Energy Use and Cost Analysis Tool*. Retrieved from http://doe2.com/DOE2/index.html

Dong, B., & Andrews, B. (2009). Sensor-based occupancy behavioral pattern recognition for energy and comfort management in intelligent buildings. *In Proceedings of building simulation*, (pp. 1444-1451).

Dong, B., & Lam, K. P. (2011). Building energy and comfort management through occupant behaviour pattern detection based on a large-scale environmental sensor network. *Journal of Building Performance Simulation, 4*(4), 359-369.

Dong, B., & Lam, K. P. (2014). A real-time model predictive control for building heating and cooling systems based on the occupancy behavior pattern detection and local weather forecasting. *Building Simulation, 7*(1), 89-106.

Dong, B., Andrews, B., Lam, K. P., Höynck, M., Zhang, R., Chiou, Y. S., & Benitez, D. (2010). An information technology enabled sustainability test-bed (ITEST) for occupancy detection through an environmental sensing network. *Energy and Buildings, 42*(7), 1038-1046.

Dong, B., Lam, K. P., & Neuman, C. (2011). Integrated building control based on occupant behavior pattern detection and local weather forecasting. *InTwelfth International IBPSA Conference* (pp. 14-17). Sydney, Australia: IBPSA.

Douc, R., Moulines, E., & Rosenthal, J. S. (2004). Quantitative bounds on convergence of time-inhomogeneous Markov chains. *Annals of Applied Probability, 14*(4), 1643-1665.

Doukas, H., Patlitzianas, K. D., Iatropoulos, K., & Psarras, J. (2007). Intelligent building energy management system using rule sets. *Building and environment, 42*(10), 3562-3569.

Duarte, C., Budwig, R., & Van Den Wymelenberg, K. (2015). Energy and demand implication of using recommended practice occupancy diversity factors compared to real occupancy data in whole building energy simulation. *Journal of Building Performance Simulation , 8*(6), 408-423.

Duarte, C., Van Den Wymelenberg, K., & Rieger, C. (2013). Revealing occupancy patterns in an office building through the use of occupancy sensor data. *Energy and buildings, 67*, 587-595.

Ekwevugbe, T., Brown, N., & Fan, D. (2012). A design model for building occupancy detection using sensor fusion. *In Digital Ecosystems Technologies (DEST), 6th IEEE International Conference* (pp. 1-6). IEEE.

Ekwevugbe, T., Brown, N., Pakka, V., & Fan, D. (2017). Improved occupancy monitoring in non-domestic buildings. *Sustainable Cities and Society, , 30*, 97-107.

Energy Information Administration (EIA). (2010). *Residential Energy Consumption Survey: Preliminary Housing Characteristics Tables.* Retrieved from U.S. Energy Information Administration (EIA): https://www.eia.gov/

EnergyPlus. (2015). *EnergyPlus Engineering Reference.* a National Laboratory of the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (NREL), Berkeley, California: LBNL Press.

EnergyPlus. (2015). *Weather Data.* Retrieved from DesignBuilder Documents: https://energyplus.net/weather

EPA. (1989). *Report to Congress on Indoor Air Quality.* Washington, DC.: U.S. Environmental Protection Agency, EPA/400/1-89/001C.

Erickson, V. L., & Cerpa, A. E. (2010). Occupancy based demand response HVAC control strategy. *In Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building* (pp. 7-12). ACM.

Erickson, V. L., Carreira-Perpiñán, M. Á., & Cerpa, A. E. (2011). OBSERVE: Occupancy-based system for efficient reduction of HVAC energy. *In Information Processing in Sensor Networks (IPSN), 10th International Conference* (pp. 258-269). Chicago, Illinois: IEEE.

Erickson, V. L., Lin, Y., Kamthe, A., Brahme, R., Surana, A., Cerpa, A. E., . . . Narayanan, S. (2009). Energy efficient building environment control strategies using real-time occupancy measurements. *In Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings* (pp. 19-24). ACM.

Ersue, M., Romascanu, D., Schoenwaelder, J., & Sehgal, A. (2015). *Management of networks with constrained devices: use cases.* (No. RFC 7548).

Escuyer, S., & Fontoynont, M. (2001). Lighting controls: a field study of office workers' reactions. *Transactions of the Illuminating Engineering Society, 33*(2), 77-94.

ESRU. (2012). *ESP-r*. Retrieved from http://www.esru.strath.ac.uk/Programs/ESP-r.htm

EU. (2018). *EU energy in figures, statistical pocketbook 2018.* Luxembourg: Imprimerie Centrale SA.

European Commission. (2003). *Indoor Air Pollution: New EU Research Reveals Higher Risks than Previously Thought.* Brussels: Press Release Database.

Fabi, V., Camisassi, V., Causone, F., Corgnati, S. P., & Andersen, R. K. (2014). Light switch behaviour: occupant behaviour stochastic models in office buildings. *In 8th windsor conference* (pp. 1-9). Windsor, UK: London: Network for Comfort and Energy Use in Buildings.

Fanger, P. (1972). *Thermal Comfort-Analysis and Applications in Environmental Engineering.* Copenhagen: Danish Technical Press.

Feng, X., Yan, D., & Hong, T. (2015). Simulation of occupancy in buildings. *Energy and Buildings, 87*, 348-359.

Fernandes, L. L., Lee, E. S., DiBartolomeo, D. L., & McNeil, A. (2014). Monitored lighting energy savings from dimmable lighting controls in The New York Times Headquarters Building. *Energy and Buildings, 68*, 498-514.

Foster, T. W., Bhatt, D. V., Hancke, G. P., & Silva, B. (2016). A Web-Based Office Climate Control System Using Wireless Sensors. *IEEE Sensors Journal, 16*(15), 6104-6113.

Gagnon, R., Gosselin, L., & Decker, S. (2018). Sensitivity analysis of energy performance and thermal comfort throughout building design process. *Energy and Buildings, 164*, 278-294.

Galasiu, A. D., & Newsham, G. R. (2009). Energy savings due to occupancy sensors and personal controls: A pilot field study. *Proceedings of Lux Europa.* Ottawa, Ontario, Canada: NSERC.

Galasiu, A. D., Newsham, G. R., Suvagau, C., & Sander, D. M. (2007). Energy saving lighting control systems for open-plan offices: a field study. *Leukos, 4*(1), 7-29.

Garg, V., & Bansal, N. K. (2000). Smart occupancy sensors to reduce energy consumption. *Energy and Buildings, 32*(1), 81-87.

Gentile, N., & Dubois, M. C. (2017). Field data and simulations to estimate the role of standby energy use of lighting control systems in individual offices. *Energy and Buildings, 155*, 390-403.

Goldstein, R., Tessier, A., & Khan, A. (2010a). Customizing the behavior of interacting occupants using personas. *In Proceedings of the National IBPSA-USA Conference*, (pp. 252-259). New York, USA.

Goldstein, R., Tessier, A., & Khan, A. (2010b). Schedule-calibrated occupant behavior simulation. *In Proceedings of the 2010 Spring Simulation Multiconference* (p. p. 180). Society for Computer Simulation International.

Goldstein, R., Tessier, A., & Khan, A. (2011). Space layout in occupant behavior simulation. *In Conference Proceedings: IBPSA-AIRAH Building Simulation Conference* (pp. 1073-1080). Sydney: 12th Conference of International Building Performance Simulation Association.

Goyal, S., Barooah, P., & Middelkoop, T. (2015). Experimental study of occupancy-based control of HVAC zones. *Applied Energy, 140*, 75-84.

Goyal, S., Ingley, H. A., & Barooah, P. (2012). Effect of various uncertainties on the performance of occupancy-based optimal control of HVAC zones. *In Decision and Control (CDC), 2012 IEEE 51st Annual Conference* (pp. 7565-7570). IEEE.

Goyal, S., Ingley, H. A., & Barooah, P. (2013). Occupancy-based zone-climate control for energy-efficient buildings: Complexity vs. performance. *Applied Energy, 106*, 209-221.

Granderson, J., & Price, P. N. (2014). Development and application of a statistical methodology to evaluate the predictive accuracy of building energy baseline models. *Energy, 66*, 981-990.

Gruber, M., Trüschel, A., & Dalenbäck, J. O. (2014). CO 2 sensors for occupancy estimations: Potential in building automation applications. *Energy and Buildings, 84*, 548-556.

Guerra-Santin, O., & Itard, L. (2012). The effect of energy performance regulations on energy consumption. *Energy Efficiency, 5*(3), 269-282.

Gunay, H. B., O'Brien, W. B.-M., & Bursill, J. (2016). Implementation of an adaptive occupancy and building learning temperature setback algorithm. *Ashrae Transactions, 122*, 179-192.

Gunay, H. B., O'Brien, W., Beausoleil-Morrison, I., Goldstein, R., Breslav, S., & Khan, A. (2014). Coupling stochastic occupant models to building performance simulation using the discrete event system specification formalism. *Journal of Building Performance Simulation, 7*(6), 457-478.

Guo, X., Tiller, D. K., Henze, G. P., & Waters, C. E. (2010). The performance of occupancy-based lighting control systems: A review. *Lighting Research & Technology, 42*(4), 415-431.

Hailemariam, E., Goldstein, R., Attar, R., & Khan, A. (2011). Real-time occupancy detection using decision trees with multiple sensor types. *In Proceedings of the 2011 Symposium on Simulation for Architecture and Urban Design* (pp. 141-148). Society for Computer Simulation International.

Haldi, F., & Robinson, D. (2008). On the behaviour and adaptation of office occupants. *Building and environment, 43*(12), 2163-2177.

Haldi, F., & Robinson, D. (2011). Modelling occupants' personal characteristics for thermal comfort prediction. *International journal of biometeorology, 55*(5), 681-694.

Han, Z., Gao, R. X., & Fan, Z. (2012). Occupancy and indoor environment quality sensing for smart buildings. *In Instrumentation and Measurement Technology Conference (I2MTC), 2012 IEEE International* (pp. 882-887). IEEE.

Harle, R. K., & Hopper, A. (2008). The potential for location-aware power management. *In Proceedings of the 10th international conference on Ubiquitous computing* (pp. 302-311). Seoul, Korea: ACM.

Harris, C., & Cahill, V. (2005). Exploiting user behaviour for context-aware power management. *In Wireless And Mobile Computing, Networking And Communications, IEEE International Conference. 4*, pp. 122-130. IEEE.

Hassan, R., Cohanim, B., De Weck, O., & Venter, G. (2005). A comparison of particle swarm optimization and the genetic algorithm. *In Proceedings of the 1st AIAA multidisciplinary design optimization specialist conference* (pp. 1-13). American Institute of Aeronautics and Astronautics.

Hightower, J., & Borriello, G. (2001). Location systems for ubiquitous computing. *Computer, 34*(8), 57-66.

Holland, J. (1975). *Adaptation in Natural and Artificial System.* Ann Arbor, MI: University of Michigan Press.

Hong, T., & Lin, H.W. (2014). Occupant behavior: impact on energy use of private offices. *In ASim 2012-1st Asia conference of International Building Performance Simulation Association*, (pp. 1-12). Shanghai, China, 11/25/12-11/27/12.

Hong, T., D'Oca, S., Turner, W. J., & Taylor-Lange, S. C. (2015a). An ontology to represent energy-related occupant behavior in buildings. Part I: Introduction to the DNAs framework. *Building and Environment, 92*, 764-777.

Hong, T., D'Oca, S., Taylor-Lange, S. C., Turner, W. J., Chen, Y., & Corgnati, S. P. (2015b). An ontology to represent energy-related occupant behavior in buildings. Part II: Implementation of the DNAS framework using an XML schema. *Building and Environment, 94*, 196-205.

Howard, B., Acha, S., Shah, N., & Polak, J. (2019). Implicit Sensing of Building Occupancy Count with Information and Communication Technology Data Sets. *Building and Environment, 157*, 297-308.

Humphreys, M. A., Rijal, H. B., & Nicol, J. F. (2013). Updating the adaptive relation between climate and comfort indoors; new insights and an extended database. *Building and Environment, 63*, 40-55.

IBM. (2016). *Embracing the Internet of Things in the new era of cognitive buildings.* United States of America: IBM Corporation.

IBM. (2017). *Energy and environment*. (IBM) Retrieved from https://www.ibm.com/ibm/green/smarter_buildings.html

IEA EBC Annex 66. (2013-2017). *International Energy Agency. Energy in buildings and communities program.* Retrieved from http://www.annex66.org/

IES. (2019). *Integrated Environmental Solutions* . Retrieved from https://www.iesve.com/

Imanishi, T., Tennekoon, R., Palensky, P., & Nishi, H. (2015). Enhanced building thermal model by using CO2 based occupancy data. *IECON 2015-41st Annual Conference of the IEEE Industrial Electronics Society* (pp. 003116-003121). IEEE.

IndustryARC. (2016). *Bluetooth Smart/Bluetooth Low Energy market: applications (consumer electronics, healthcare, sports & fitness, retail, automotive, security); by technology [discrete modules, integrated modules (single & dual mode)]-forecast (2017-2022).* Retrieved from http://industryarc.com

Ioannou, A., & Itard, L. C. (2015). Energy performance and comfort in residential buildings: Sensitivity for building parameters and occupancy. *Energy and Buildings, 92*, 216-233.

Jain, S., & Madamopoulos, N. (2016). Ahorrar: Indoor Occupancy Counting to Enable Smart Energy Efficient Office Buildings. *In Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom)(BDCloud-SocialCom-SustainCom)* (pp. 469-476). IEEE.

Javed, A., Larijani, H., Ahmadinia, A., Emmanuel, R., Mannion, M., & Gibson, D. (2016). Design and implementation of a cloud enabled random neural network-based decentralized smart controller with intelligent sensor nodes for HVAC. *IEEE Internet of Things, 4*(2), 393-403.

Jazizadeh, F., & Becerik-Gerber, B. (2012). A novel method for non intrusive load monitoring of lighting systems in commercial buildings. *Computing in Civil Engineering*, 523-530.

Jennings, J. D., Rubinstein, F. M., DiBartolomeo, D., & Blanc, S. L. (2000). Comparison of control options in private offices in an advanced lighting controls testbed. *Journal of the Illuminating Engineering Society, 29*(2), 39-60.

Jennings, J., Colak, N., & Rubinstein, F. (2002). Occupancy and time-based lighting controls in open offices. *Journal of the Illuminating Engineering Society, 31*(2), 86-100.

Jiefan, G., Peng, X., Zhihong, P., Yongbao, C., Ying, J., & Zhe, C. (2018). Extracting typical occupancy data of different buildings from mobile positioning data. *Energy and Buildings, 180*, 135-145.

Jin, M. J., & Spanos, C. (2017). Virtual occupancy sensing: Using smart meters to indicate your presence. *IEEE Transactions on Mobile Computing* (pp. 1-14). IEEE. doi:DOI 10.1109/TMC.2017.2684806

Jin, M., Jia, R., Kang, Z., Konstantakopoulos, I. C., & Spanos, C. J. (2014). Presencesense: Zero-training algorithm for individual presence detection based on power monitoring. *In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings* (pp. 1-10). ACM.

Justel, A., Peña, D., & Zamar, R. (1997). A multivariate Kolmogorov-Smirnov test of goodness of fit. *Statistics & Probability Letters, 35*(3), 251-259.

Kamthe, A., Jiang, L., Dudys, M., & Cerpa, A. (2009). Scopes: Smart cameras object position estimation system. In *Wireless Sensor Networks* (pp. 279-295). Berlin, Heidelberg: In European Conference on Wireless Sensor Networks, Springer.

Karjalainen, S. (2007). Gender differences in thermal comfort and use of thermostats in everyday thermal environments. *Building and environment, 42*(4), 1594-1603.

Kavulya, G., & Becerik-Gerber, B. (2012). Understanding the influence of occupant behavior on energy consumption patterns in commercial buildings. *Computing in Civil Engineering*, 569-576.

Khan, A., & Hornbæk, K. (2011). Big data from the built environment. *Proceedings of the 2nd international workshop on Research in the large* (pp. 29-32). Beijing, China: ACM.

Khan, A., Nicholson, J., Mellor, S., Jackson, D., Ladha, K., Ladha, C., . . . Plötz, T. (2014). Occupancy monitoring using environmental & context sensors and a hierarchical analysis framework . *BuildSys'14* (pp. 90-99). Memphis, TN, USA: ACM.

Kim, J., Bauman, F., Raftery, P., Arens, E., Zhang, H., Fierro, G., & ... & Culler, D. (2019). Occupant comfort and behavior: High-resolution data from a 6-month field study of personal comfort systems with 37 real office workers. *Building and Environment, 148*, 348-360.

Kim, Y. S., & Srebric, J. (2015). Improvement of building energy simulation accuracy with occupancy schedules derived from hourly building electricity consumption. *Ashrae Transactions, 121*(1), 353-361.

Krishnamachari, B., Estrin, D., & Wicker, S. B. (2002). The Impact of Data Aggregation in Wireless Sensor Networks. *Proceedings of the 22 nd International Conference on Distributed Computing Systems Workshops (ICDCSW'02). 578*, pp. 1-8. ICDCS workshops.

Kuo, T. C., Chan, Y. C., & Chen, A. Y. (2017). An Occupant-Centered Integrated Lighting and Shading Control for Energy Saving and Individual Preferences. *In Computing in Civil Engineering 2017* (pp. 207-214). Seattle, USA: IWCCE.

Labeodan, T., Aduda, K., Zeiler, W., & Hoving, F. (2016). Experimental evaluation of the performance of chair sensors in an office space for occupancy detection and occupancy-driven control. *Energy and Buildings, 111*, 195-206.

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

Lam, K. P., Höynck, M., Dong, B., Andrews, B., Chiou, Y. S., Zhang, R., . . . Choi, J. (2009). Occupancy detection through an extensive environmental sensor network in an open-plan office building. , 145,. *IBPSA Building Simulation. 145*, pp. 1452-1459. Glasgow, Scotland: Eleventh International IBPSA Conference.

Langevin, J., Wen, J., & Gurian, P. L. (2013). Modeling thermal comfort holistically: Bayesian estimation of thermal sensation, acceptability, and preference distributions for office building occupants. *Building and Environment, 69*, 206-226.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436-444.

Lee, D., Sim, A., Choi, J., & Wu, K. (2016). Novel data reduction based on statistical similarity. *Proceedings of the 28th International Conference on Scientific and Statistical Database Management* (pp. 21-33). ACM.

Lejlic, E. (2017). *The Advantages of Machine Learning.* Lumagate.

Li, N., Calis, G., & Becerik-Gerber, B. (2012). Measuring and monitoring occupancy with an RFID based system for demand-driven HVAC operations. *Automation in construction, 24*, 89-99.

Li, S., Deng, K., & Zhou, M. (2016). Sensitivity analysis for building energy simulation model calibration via algorithmic differentiation. *IEEE Transactions on Automation Science and Engineering , 14*(2), 905-914.

Liang, X., Hong, T., & Shen, G. Q. (2016). Improving the accuracy of energy baseline models for commercial buildings with occupancy data. *Applied energy, 179*, 247-260.

Liao, C., & Barooah, P. (2010). An integrated approach to occupancy modeling and estimation in commercial buildings. *In Proceedings of the 2010 American Control Conference* (pp. 3130-3135). IEEE.

Lilis, G., Conus, G., Asadi, N., & Kayal, M. (2017). Towards the next generation of intelligent building: An assessment study of current automation and future IoT based systems with a proposal for transitional design. *Sustainable cities and society, 28*, 473-481.

Lim, B., Hijazi, H., Thiébaux, S., & van den Briel, M. (2016). Online HVAC-Aware Occupancy Scheduling with Adaptive Temperature Control. *In International Conference on Principles and Practice of Constraint Programming* (pp. 683-700). Springer International Publishing.

Lim, G. H., Keumala, N., & Ghafar, N. A. (2017). Energy saving potential and visual comfort of task light usage for offices in Malaysia. *Energy and Buildings, 147*, 166-175.

Liu, Z. (2017). *Simulation of Local Climate Control in Shared Offices Based on Occupants Locations and Preferences.* Montreal: Master Thesis, Concordia University.

Liu, Z., Salimi, S., & Hammad, A. (2016). Simulation of HVAC Local Control Based on Occupants Locations and Preferences. *33rd International Symposium on Automation and Robotics in Construction* . Auburn, Alabama, USA: ISARC.

Lo, L. J., & Novoselac, A. (2010). Localized air-conditioning with occupancy control in an open office. *Energy and Buildings, 42*(7), 1120-1128.

Mackensen, E., Lai, M., & Wendt, T. M. (2012). Performance analysis of an Bluetooth Low Energy sensor system. *In Wireless Systems (IDAACS-SWS), IEEE 1st International Symposium* (pp. 62-66). IEEE.

Mahdavi, A., Mohammadi, A., Kabir, E., & Lambeva, L. (2008). Occupants' operation of lighting and shading systems in office buildings. *Journal of building performance simulation, 1*(1), 57-65.

Mahdavi, A., Mohammadi, A., Kabir, E., & Lambeva, L. (2008). Shading and lighting operation in office buildings in Austria: A study of user control behavior. *In Building simulation, 1*(2), 111-117.

Majcen, D., Itard, L. C., & Visscher, H. (2013a). Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications. *Energy policy, 54*, 125-136.

Majcen, D., Itard, L., & Visscher, H. (2013b). Actual and theoretical gas consumption in Dutch dwellings: What causes the differences? *Energy Policy, 61*, 460-471.

Majumdar, A., Setter, J. L., Dobbs, J. R., Hencey, B. M., & Albonesi, D. H. (2014). Energy-comfort optimization using discomfort history and probabilistic occupancy prediction. *In Green Computing Conference (IGCC)* (pp. 1-10). IEEE.

Maniccia, D., Tweed, A., Bierman, A., & Von Neida, B. (2001). The effects of changing occupancy sensor time-out setting on energy savings, lamp cycling and maintenance costs. *Journal of the Illuminating Engineering Society, 30*(2), 97-110.

Manzoor, F., Linton, D., & Loughlin, M. (2012). Occupancy monitoring using passive RFID technology for efficient building lighting control. *In RFID Technology (EURASIP RFID), 2012 Fourth International EURASIP Workshop* (pp. 83-88). IEEE.

Mashuk, M. S., Pinchin, J., Siebers, P. O., & Moore, T. (2018). A smart phone based multi-floor indoor positioning system for occupancy detection. *Position, Location and Navigation Symposium (PLANS)* (pp. 216-227). IEEE.

Masoso, O. T., & Grobler, L. J. (2010). The dark side of occupants' behaviour on building energy use.

Massart, D. L., & Smeyers-verbeke, A. J. (2005). Practical data handling visual presentation of data by means of box plots. *LCGC Europe, 18*(4), 215-2018.

Massey Jr, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association, 46*(253), 68-78.

McCall, J. (2005). Genetic algorithms for modelling and optimisation. *Journal of Computational and Applied Mathematics, 184*(1), 205-222.

Mellouk, L., Aaroud, A., Benhaddou, D., Zine-Dine, K., & Boulmalf, M. (2015). Overview of mathematical methods for energy management optimization in smart grids. *In Renewable and Sustainable Energy Conference (IRSEC), 2015 3rd International* (pp. 1-5). IEEE.

Meyn, S., Surana, A., Lin, Y., Oggianu, S. M., Narayanan, S., & Frewen, T. A. (2009). A sensor-utility-network method for estimation of occupancy in buildings . *In Decision and Control, held jointly with the 28th Chinese Control Confere* (pp. 1494-1500). Proceedings of the 48th IEEE Conference.

Milenkovic, M., & Amft, O. (2013). Recognizing energy-related activities using sensors commonly installed in office buildings. *Procedia Computer Science, 19*, 669-677.

Mohammadmoradi, H., Yin, S., & Gnawali, O. (2017). Room occupancy estimation through WiFi, UWB, and light sensors mounted on doorways . *Proceedings of the 2017 International Conference on Smart Digital Environment* (pp. 27-34). ACM.

Moreno, M. V., Dufour, L., Skarmeta, A. F., Jara, A. J., Genoud, D., Ladevie, B., & Bezian, J. J. (2016). Big data: the key to energy efficiency in smart buildings. *Soft Computing, 20*(5), 1749-1762.

Mustapa, M. S., Zaki, S. A., Rijal, H. B., Hagishima, A., & Ali, M. S. (2016). Thermal comfort and occupant adaptive behaviour in Japanese university buildings with free running and cooling mode offices during summer. *Building and Environment, 105*, 332-342.

Nagarathinam, S., Doddi, H., Vasan, A., Sarangan, V., Ramakrishna, P. V., & Sivasubramaniam, A. (2017). Energy efficient thermal comfort in open-plan office buildings. *Energy and Buildings, 139*, 476-486.

Nagy, Z., Yong, F. Y., & Schlueter, A. (2016). Occupant centered lighting control: A user study on balancing comfort, acceptance, and energy consumption. *Energy and Buildings, 126*, 310-322.

Nagy, Z., Yong, F. Y., Frei, M., & Schlueter, A. (2015). Occupant centered lighting control for comfort and energy efficient building operation. *Energy and Buildings, 94*, 100-108.

Nasir, N., Palani, K., Chugh, A., Prakash, V. C., Arote, U., Krishnan, A. P., & Ramamritham, K. (2015). Fusing sensors for occupancy sensing in smart buildings. *In International Conference on Distributed Computing and Internet Technology* (pp. 73-92). Springer, Cham.

Nesa, N., & Banerjee, I. (2017). IoT-based sensor data fusion for occupancy sensing using Dempster–Shafer evidence theory for smart buildings. *IEEE Internet of Things Journal, 4*(5), 1563-1570.

Newsham, G. R., & Birt, B. J. (2010). Building-level occupancy data to improve ARIMA-based electricity use forecasts. *In Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building* (pp. 13-18). ACM.

Newsham, G. R., Xue, H., Arsenault, C., Valdes, J. J., Burns, G. J., Scarlett, E., & ... & Shen, W. (2017). Testing the accuracy of low-cost data streams for determining single-person office occupancy and their use for energy reduction of building services. *Energy and Buildings, 135*, 137-147.

Nguyen, T. A., & Aiello, M. (2012). Beyond Indoor Presence Monitoring with Simple Sensors. *PECCS*, 5-14.

Nguyen, T. A., & Aiello, M. (2013). Energy intelligent buildings based on user activity: A survey. *Energy and buildings, 56*, 244-257.

NLPIP, N. L. (1998). *Occupancy Sensors: Motion Sensors for Lighting Control.* Troy (NY): Lighting Research Center.

NRCAN. (2012). *Improving Energy Performance in Canada.* Minister of Natural Resources Canada, Government of Canada.

NRCAN. (2016). *Improving Energy Performance in Canada.* Natural Resources Canada.

NRCAN. (2018). *Energy Efficiency Trends Analysis Tables.* Ottawa: Statistics Canada.

Oldewurtel, F., Sturzenegger, D., & Morari, M. (2013). Importance of occupancy information for building climate control. *Applied energy, 101*(1), 521-532.

Ott, R. L., & Longnecker, M. T. (2015). *An introduction to statistical methods and data analysis.* Nelson Education.

Page, J., Robinson, D., Morel, N., & Scartezzini, J. L. (2008). A generalised stochastic model for the simulation of occupant presence. *Energy and buildings, 40*(2), 83-98.

Panda, S., & Padhy, N. P. (2008). Comparison of particle swarm optimization and genetic algorithm for FACTS-based controller design. *Applied soft computing, 8*(4), 1418-1427.

Pandharipande, A., & Caicedo, D. (2011). Daylight integrated illumination control of LED systems based on enhanced presence sensing. *Energy and Buildings, 43*(4), 944-950.

Pandharipande, A., & Caicedo, D. (2015). Smart indoor lighting systems with luminaire-based sensing: A review of lighting control approaches. *Energy and Buildings, 104*, 369-377.

Pang, Z., Xu, P., O'Neill, Z., Gu, J., Qiu, S., Lu, X., & Li, X. (2018). Application of mobile positioning occupancy data for building energy simulation: An engineering case study. *Building and Environment, 141*, 1-15.

Papantoniou, S., Kolokotsa, D., & Kalaitzakis, K. (2015). Building optimization and control algorithms implemented in existing BEMS using a web based energy management and control system. *Energy and Buildings, 98*, 45-55.

Pasini, D., Ventura, S. M., Rinaldi, S., Bellagente, P., Flammini, A., & Ciribini, A. L. (2016). Exploiting Internet of Things and building information modeling framework for management of cognitive buildings. *In Smart Cities Conference (ISC2)* (pp. 1-6). IEEE.

Patel, K. K., Patel, S. M., & Scolar, P. G. (2016). Internet of Things-IOT: Definition, Characteristics, Architecture, Enabling Technologies, Application & Future Challenges. *International Journal of Engineering Science*, 6122-6131.

Pavlovas, V. (2004). Demand controlled ventilation: A case study for existing Swedish multifamily buildings. *Energy and buildings, 36*(10), 1029-1034.

Peng, Y., Rysanek, A., Nagy, Z., & Schlüter, A. (2017). Occupancy learning-based demand-driven cooling control for office spaces. *Building and Environment, 122*, 145-160.

Peruffo, A., Pandharipande, A., Caicedo, D., & Schenato, L. (2015). Lighting control with distributed wireless sensing and actuation for daylight and occupancy adaptation. *Energy and Buildings, 97*, 13-20.

Pfafferott, J., & Herkel, S. (2007). Statistical simulation of user behaviour in low-energy office buildings. *Solar Energy, 81*(5), 676-682.

Pingel, J. (2017). *Introduction to Deep Learning: Machine Learning vs. Deep Learning.* MathWorks.

Piselli, C., & Pisello, A. L. (2019). Occupant behavior long-term continuous monitoring integrated to prediction models: Impact on office building energy performance. *Energy, 176*, 667-681.

Pradhan, A., Ergen, E., & Akinci, B. (2009). Technological assessment of radio frequency identification technology for indoor localization. *Journal of Computing in Civil Engineering, 23*(4), 230-238.

Purdon, S., Kusy, B., Jurdak, R., & Challen, G. (2013). Model-free HVAC control using occupant feedback. *In Local Computer Networks Workshops (LCN Workshops)* (pp. 84-92). IEEE.

Quuppa. (2017). *One for all*. Retrieved from Quuppa Do More With Location: http://quuppa.com/applications/

Quuppa Intelligent Locating System™. (2016). *Unique Technology*. Retrieved from Quuppa Do More With Location: http://quuppa.com/technology/

Raji, B., Tenpierik, M. J., & van den Dobbelsteen, A. (2017). Early-stage design considerations for the energy-efficiency of high-rise office buildings. *Sustainability, 9*(4), 623.

Richardson, I., Thomson, M., & Infield, D. (2008). A high-resolution domestic building occupancy model for energy demand simulations. *Energy and buildings, 40*(8), 1560-1566.

Roselli, L., Mariotti, C., Mezzanotte, P., Alimenti, F., Orecchini, G., Virili, M., & Carvalho, N. B. (2015). Review of the present technologies concurrently contributing to the implementation of the Internet of Things (IoT) paradigm: RFID, Green Electronics, WPT and Energy Harvesting. *IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNet)* (pp. 1-3). IEEE.

Rossi, M., Pandharipande, A., Caicedo, D., Schenato, L., & Cenedese, A. (2015). Personal lighting control with occupancy and daylight adaptation. *Energy and Buildings, 105*, 263-272.

Rubinstein, F., & Enscoe, A. (2010). Saving energy with highly-controlled lighting in an open-plan office. *Leukos, 7(1),, 7*(1), 21-36.

Salimi, S., & Hammad, A. (2018). Critical Review and Research Roadmap of Office Building Energy Management Based on Occupancy Monitoring. *Energy and Buildings, 182*, 214-241. doi:10.1016/j.enbuild.2018.10.007

Salimi, S., Liu, Z., & Hammad, A. (2017). Simulation-based Optimization of Energy Consumption and Discomfort in Multi-Occupied Offices Considering Occupants Locations and Preferences. *Proceedings of the 15th IBPSA Conference.* San Francisco, CA, USA.

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.

Sangogboye, F. C., Arendt, K., Jradi, M., Veje, C., Kjærgaard, M. B., & Jørgensen, B. N. (2018). The impact of occupancy resolution on the accuracy of building energy performance simulation. *Proceedings of the 5th Conference on Systems for Built Environments* (pp. 103-106). Shenzen, China: ACM.

Santucci, G. (2010). The Internet of Things: Between the Revolution of the Internet and the Metamorphosis of Objects. *Vision and Challenges for Realising the Internet of Things*, 11-24.

Sehar, F., Pipattanasomporn, M., & Rahman, S. (2016). A peak-load reduction computing tool sensitive to commercial building environmental preferences. *Applied energy, 161*, 279-289.

Sehar, F., Pipattanasomporn, M., & Rahman, S. (2017). Integrated automation for optimal demand management in commercial buildings considering occupant comfort. *Sustainable Cities and Society, 28*, 16-29.

Serfozo, R. (2009). *Basics of Applied Stochastic Processes, Probability and its Applications.* Verlag Berlin Heidelberg: Springer.

Shaikh, P. H., Nor, N. B., Nallagownden, P., & Elamvazuthi, I. (2016). Intelligent multi-objective optimization for building energy and comfort management. *Journal of King Saud University-Engineering Sciences*, 1-10.

Shen, W., Newsham, G., & Gunay, B. (2017). Leveraging existing occupancy-related data for optimal control of commercial office buildings: A review. *Advanced Engineering Informatics*, 1-13.

Shi, Y. (2001). Particle swarm optimization: developments, applications and resources. *In evolutionary computation* (pp. 81-86). IEEE.

Shih, H. C. (2014). A robust occupancy detection and tracking algorithm for the automatic monitoring and commissioning of a building. *Energy and Buildings, 77*, 270-280.

Si, W., Pan, X., & Ogai, H. (2017). Real-time Daylight Modeling Method for Lighting Systems Based on RBFNN. *In Proceedings of the 9th International Conference on Computer and Automation Engineering* (pp. 318-325). ACM.

Singhvi, V., Krause, A., Guestrin, C., Garrett Jr, J. H., & Matthews, H. S. (2005). Intelligent light control using sensor networks. In Proceedings of the 3rd international conference on Embedded networked sensor systems. (pp. 218-229). ACM.

Soebarto, V. I., & Williamson, T. J. (2001). Multi-criteria assessment of building performance: theory and implementation . *Building and environment, 36*(6), 681-690.

Soltani, M. M., Motamedi, A., & Hammad, A. (2015). Enhancing Cluster-based RFID Tag Localization using artificial neural networks and virtual reference tags. *Automation in Construction, 54*, 93-105.

Somayajulu, S. (2014, 10 30). *Automatic Vs. Smart Buildings*. Retrieved from XCHANGING BLOG : http://www.xchanging.com/blog/2014/10/30/automatic-vs-smart-buildings

Srinivas, N., & Deb, K. (1994). Muiltiobjective optimization using nondominated sorting in genetic algorithms. . *Evolutionary computation, 2*(3), 221-248.

Srivastava, T. (2015). *Difference between Machine Learning & Statistical Modeling.* Analytics Vidhya - Learn everything about Analytics.

Sun, J., Feng, B., & Xu, W. (2004). Particle swarm optimization with particles having quantum behavior. *In Evolutionary Computation, CEC2004* (pp. 325-331). IEEE.

Sun, K., Yan, D., Hong, T., & Guo, S. (2014). Stochastic modeling of overtime occupancy and its application in building energy simulation and calibration. *Building and Environment, 79*, 1-12.

SUSRIS. (2013, July 25). *International Energy Outlook 2013*. Retrieved from Saudi-US Relations Inforamtion Service: http://susris.com/2013/07/25/international-energy-outlook-2013/

Tabak, V. (2008). *User simulation of space utilisation.* Eindhoven University Press.

Tabak, V., & de Vries, B. (2010). Methods for the prediction of intermediate activities by office occupants. *Building and Environment, 45*(6), 1366-1372.

Tahmasebi, F., & Mahdavi, A. (2015). A systematic assessment of the sensitivity of building performance simulation results with regard to Occupancy-related input assumptions. *14th International Conference of the International Building Performance Simulation Association (BS 2015)*, (pp. 1397-1403). Hyderabad, India.

Tahmasebi, F., & Mahdavi, A. (2018). On the utility of occupants' behavioural diversity information for building performance simulation: An exploratory case study. *Energy and Buildings, 176*, 380-389.

Tanimoto, J., Hagishima, A., & Sagara, H. (2008). Validation of methodology for utility demand prediction considering actual variations in inhabitant behaviour schedules. *Journal of Building Performance Simulation, 1(1),, 1*(1), 31-42.

The Crystal. (2016). *One of the World's Most Sustainable Buildings*. (Siemens plc) Retrieved from https://www.thecrystal.org/

Thomas, V. C. (2018). *Heat gains and losses: roofs and walls.*

Tian, Z., Yang, J., Lei, Y. P., & Yang, L. (2019). Sensitivity Analysis of Infiltration Rates Impact on Office Building Energy Performance. *IOP Conference Series: Earth and Environmental Science. 238*, p. 012019. IOP Publishing.

Tiller, D. K., Guo, X., Henze, G. P., & Waters, C. E. (2009). The application of sensor networks to lighting control. *Leukos, 5(4),, 5*(4), 313-325.

TRNSYS. (2013). *A TRaNsient SYstem simulation program, Version 17.1.* Madison, Wisconsin: Copyright of the Board of Regents of the University of Wisconsin.

Tsanas, A., & Xifara, A. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings, 49*, 560-567.

U.S. EIA. (2019). *Annual Energy Outlook 2019 with projections to 2050.* Washington, DC: U.S. Department of Energy.

van de Meugheuvel, N., Pandharipande, A., Caicedo, D., & Van Den Hof, P. P. (2014). Distributed lighting control with daylight and occupancy adaptation. *Energy and Buildings, 75*, 321-329.

Virote, J., & Neves-Silva, R. (2012). Stochastic models for building energy prediction based on occupant behavior assessment. *Energy and Buildings, 53*, 183-193.

Von Neida, B., Manicria, D., & Tweed, A. (2001). An analysis of the energy and cost savings potential of occupancy sensors for commercial lighting systems. *Journal of the Illuminating Engineering Society, 30*(2), 111-125.

Wang, C., Yan, D., & Jiang, Y. (2011). A novel approach for building occupancy simulation. *In Building simulation , 4*(2), 149-167.

Wang, D., Federspiel, C. C., & Rubinstein, F. (2005). Modeling occupancy in single person offices. *Energy and buildings, 37*(2), 121-126.

Wang, J., Zhai, Z. J., Jing, Y., Zhang, X., & Zhang, C. (2011). Sensitivity analysis of optimal model on building cooling heating and power system. *Applied Energy, 88*(12), 5143-5152.

Wang, R. (2016). An Improved Nondominated Sorting Genetic Algorithm for Multiobjective Problem. *Mathematical Problems in Engineering* (pp. 1-7). Hindawi Publishing Corporation.

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, W., Chen, J., Huang, G., & Lu, Y. (2017). Energy efficient HVAC control for an IPS-enabled large space in commercial buildings through dynamic spatial occupancy distribution. *Applied Energy*, 1-19.

Wang, W., Wang, J., Chen, J., Huang, G., & Guo, X. (2018). Multi-zone outdoor air coordination through Wi-Fi probe-based occupancy sensing. *Energy and Buildings, 159*, 495-507.

Wang, Y., & Shao, L. (2017a). Understanding occupancy and user behaviour through Wi-Fi-based indoor positioning. *Building Research & Information*, 1-13.

Wang, Y., & Shao, L. (2017b). Understanding occupancy pattern and improving building energy efficiency through Wi-Fi based indoor positioning. *Building and Environment, 114*, 106-117.

Wang, Z., & Ding, Y. (2015). An occupant-based energy consumption prediction model for office equipment. *Energy and Buildings, 109*, 12-22.

Ward, A. M. (2007). In-Building Location Systems. *Proceedings of the the Institution of Engineering and Technology Seminar on Location Technologies*, (pp. 1-18). London, U.K.

Wei, S., Buswell, R., & Loveday, D. (2010). Probabilistic modelling of human adaptive behaviour in non-airconditioned buildings. *Adapting to Change: New Thinking on Comfort* (pp. 1-17). Cumberland Lodge, Windsor, UK: Network for Comfort and Energy Use in Buildings.

Wei, S., Yong, J., Ng, B., Tindall, J., Lu, Q., & Du, H. (2018). Occupant Adaptive Behaviour: an Effective Method towards Energy Efficient Buildings. *CIBSE Technical Symposium.* London, UK.

Wen, Y. J., & Agogino, A. M. (2008). Wireless networked lighting systems for optimizing energy savings and user satisfaction. *In Wireless Hive Networks Conference, 2008. WHNC 2008. IEEE* (pp. 1-7). IEEE.

Wen, Y. J., & Agogino, A. M. (2011). Control of wireless-networked lighting in open-plan offices. *Lighting Research & Technology, 43*(2), 235-248.

West, S. R., Ward, J. K., & Wall, J. (2014). Trial results from a model predictive control and optimisation system for commercial building HVAC. *Energy and Buildings, 72*, 271-279.

Wickham, H., & Stryjewski, L. (2011). 40 years of boxplots. *Statistician*, 1-17.

Wong, J. K., Li, H., & Wang, S. W. (2005). Intelligent building research: a review. *Automation in Construction, 14*(1), 143-159.

Yamaguchi, Y., Shimoda, Y., & Mizuno, M. (2003). Development of district energy system simulation model based on detailed energy demand model. *In Proceeding of Eighth International IBPSA Conference*, (pp. 1443-1450). Eindhoven, Netherlands.

Yan, D., O'Brien, W., Hong, T., Feng, X., Gunay, H. B., Tahmasebi, F., & Mahdavi, A. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings, 107*, 264-278.

Yang, J., Zou, H., Jiang, H., & Xie, L. (2018). Device-free occupant activity sensing using wifi-enabled iot devices for smart homes. *IEEE Internet of Things Journal, 5*(5), 3991-4002.

Yang, Z., & Becerik-Gerber, B. (2014). Modeling personalized occupancy profiles for representing long term patterns by using ambient context. *Building and Environment, 78*, 23-35.

Yang, Z., Li, N., Becerik-Gerber, B., & Orosz, M. (2012). A multi-sensor based occupancy estimation model for supporting demand driven HVAC operations. *In Proceedings of the 2012 Symposium on Simulation for Architecture and Urban Design.* Society for Computer Simulation International.

Yip, S., Athienitis, A., & Lee, B. (2019). Sensitivity analysis of building form and BIPVT energy performance for net-zero energy early-design stage consideration. *IOP Conference Series: Earth and Environmental Science*, *238*, p. 012065.

Yu, T. (2010). Modeling occupancy behavior for energy efficiency and occupants comfort management in intelligent buildings. *In Machine Learning and Applications (ICMLA), Ninth International Conference* (pp. 726-731). IEEE.

Yu, Z., Fung, B. C., Haghighat, F., Yoshino, H., & Morofsky, E. (2011). A systematic procedure to study the influence of occupant behavior on building energy consumption. *Energy and Buildings, 43*(6), 1409-1417.

Yun, G. Y., Kim, H., & Kim, J. T. (2012). Effects of occupancy and lighting use patterns on lighting energy consumption. *Energy and Buildings, 46*, 152-158.

Zhang, B., Liu, Y., Rai, R., & Krovi, V. (2017). Invariant probabilistic sensitivity analysis for building energy models. *Journal of Building Performance Simulation, 10*(4), 392-405.

Zhang, Y. (2009). Parallel EnergyPlus and the development of a parametric analysis tool. *11th Conference of International Building Performance Association (IBPSA)*, (pp. 1382-1388). Glasgow, UK.

Zhang, Y. (2012). Use jEPlus as an efficient building design optimisation tool. *CIBSE ASHRAE Technical Symposium.* London. Retrieved from https://www.jeplus.org/wiki/doku.php?id=docs:jeplus_ea:start

Zhao, R., Sun, S., & Ding, R. (2004). Conditioning strategies of indoor thermal environment in warm climates. *Energy and buildings, 36*(12), 1281-1286.

Zhao, Y., Zeiler, W., Boxem, G., & Labeodan, T. (2015). Virtual occupancy sensors for real-time occupancy information in buildings. *Building and Environment, 93*, 9-20.

Zhao, Z., Kuendig, S., Carrera, J., Carron, B., Braun, T., & Rolim, J. (2017). Indoor location for smart environments with wireless sensor and actuator networks. *IEEE 42nd Conference on Local Computer Networks (LCN)* (pp. 535-538). IEEE.

Zhao, Z., Min, G., Gao, W., Wu, Y., Duan, H., & Ni, Q. (2018). Deploying Edge Computing Nodes for Large-scale IoT: A Diversity Aware Approach. *IEEE Internet of Things Journal*, 1-8.

Zhen, Z. N., Jia, Q. S., Song, C., & Guan, X. (2008). An indoor localization algorithm for lighting control using RFID. *In Energy 2030 Conference. ENERGY 2008.* (pp. 1-6). IEEE.

Zhu, P., Gilbride, M., Yan, D., Sun, H., & Meek, C. (2017). Lighting energy consumption in ultra-low energy buildings: Using a simulation and measurement methodology to model occupant behavior and lighting controls. *Building Simulation, 10*(6), 799-810.

Zhu, Y., Hinds, W. C., Krudysz, M., Kuhn, T., Froines, J., & Sioutas, C. (2005). Penetration of freeway ultrafine particles into indoor environments. *Journal of Aerosol Science, 36*(3), 303-322.

Zou, H., Zhou, Y., Yang, J., & Spanos, C. J. (2018). Device-free occupancy detection and crowd counting in smart buildings with WiFi-enabled IoT. *Energy and Buildings, 174*, 309-322.

APPENDICES

Appendix A – MATLAB Code of Occupancy Detection Model

```
function output = IFM_Val7_1Min_Shide_PredictionVsTestData(~)
 % x and x1 are inputs:
 %inputs are day of the week, occupants name, and location of the occupant:
x = {'Shide-S-P-AllInstances-', 'Wed'};
x1 = {'Shide-S-P-AllInstances-', 'Wed', '-TestData'};

d=x{2};
name     = sprintf('%s%s', x{1}, x{2});
name1     = sprintf('%s%s%s', x1{1}, x1{2}, x1{3});
full_filename = fullfile(d,name);
full_filename1 = fullfile(d,name1);
P = xlsread(full_filename, 'PWholeDay-WholeYear');
Loc = xlsread(full_filename1, 'SWholeDay-WholeYear');
format long g
ST=P(1,2);
ET=P(end,2);
ST=datetime(ST,'ConvertFrom','datenum');
ST = datetime(ST,'Format','HH:mm');
ST = dateshift(ST, 'start', 'minute', 'nearest');
NearestHrST = dateshift(ST, 'start', 'hour', 'nearest');
MinST = minute(ST);
HrST=hour(NearestHrST);
if MinST >= 30

else
    HrST=HrST+1;
end

ET=datetime(ET,'ConvertFrom','datenum');
ET = datetime(ET,'Format','HH:mm');
ET = dateshift(ET, 'start', 'minute', 'nearest');
MinET = minute(ET);
HrET=hour(ET);

    HrDiff=HrET-HrST;
    MinDiff=MinET+(60-MinST);

    PT=(HrDiff*60)+MinDiff;
    PMatrix_Index=(PT)+1;
T(1)=ST;
Time = zeros(1, PMatrix_Index);
y = zeros(1, PMatrix_Index);

ncols= zeros(1,PMatrix_Index);
L=zeros(1,PMatrix_Index);
mat_max_index=cell(PMatrix_Index,1);
i=1;
Time(i) = rem(datenum(T(1)),1);
    zz=Loc(i, :);
    zz(:,1) = [];
        mat_max_index{i,:}=find(zz == max(zz));
        [~,ncols(i)] = cellfun(@size,mat_max_index(i,1));
            if ncols(i) > 1
```

```
                L(i)=5;
            else
                L(i)=mat_max_index{i};
            end
y (1)=L(i);
    [a,~]=find(abs(P(:,2)-Time(i))<0.00001,5);
    PR=P(a, :);
    A=PR(L(i),:);
    kk=find(A(1,:)>=0,5,'last');
    Pr=A(kk);
    [~, n] = size(Pr);
    if Pr == zeros(1,n)
        Predicted_L=1;
        A=PR(Predicted_L,:);
        kk=find(A(1,:)>=0,5,'last');
        Pr=A(kk);
          if Pr == zeros(1,n)
             while true
                 Predicted_L = Predicted_L+1;
                 A=PR(Predicted_L,:);
                 kk=find(A(1,:)>=0,5,'last');
                 Pr=A(kk);
                    if any(Pr > zeros(1,n))
                        break;
                    end
             end
          end
        pd(:,1)=Pr(:,1);
        for j = 2 : n
            pd(:,j)=Pr(:,j)+pd(:,j-1);
        end
        C=rand;
        D=pd-C;
        E=find(D>0,1);
        y(i+1)=E;
        T(2) = T(1) + minutes(1);
    else
        pd(:,1)=Pr(:,1);
        for j = 2 : n
            pd(:,j)=Pr(:,j)+pd(:,j-1);
        end

        C=rand;
        D=pd-C;
        E=find(D>0,1);
        y(i+1)=E;

        T(2) = T(1) + minutes(1);
    end
for i = 2 : (PMatrix_Index-1)

    Time(i) = rem(datenum(T(i)),1);
    zz=Loc(i, :);
    zz(:,1) = [];
    mat_max_index{i,:}=find(zz == max(zz));
    [~,ncols(i)] = cellfun(@size,mat_max_index(i,1));
```

```
    if ncols(i) > 1
        if i==1
            L(i)=5;
        else
            L(i)=L(i-1);
        end
    else
        L(i)=mat_max_index{i};
    end
    [a,~]=find(abs(P(:,2)-Time(i))<0.00001,5);

    PR=P(a, :);
    A=PR(L(i),:);
    kk=find(A(1,:)>=0,5,'last');
    Pr=A(kk);
    [~, n] = size(Pr);
    if Pr == zeros(1,n)
        Predicted_L=1;
        A=PR(Predicted_L,:);
        kk=find(A(1,:)>=0,5,'last');
        Pr=A(kk);
          if Pr == zeros(1,n)
             while true
                 Predicted_L = Predicted_L+1;
                 A=PR(Predicted_L,:);
                 kk=find(A(1,:)>=0,5,'last');
                 Pr=A(kk);
                    if any(Pr > zeros(1,n))
                        break;
                    end
              end
          end
        pd(:,1)=Pr(:,1);
        for j = 2 : n
            pd(:,j)=Pr(:,j)+pd(:,j-1);
        end
        C=rand;
        D=pd-C;
        E=find(D>0,1);
        y(i+1)=E;
        T(i+1) = T(i) + minutes(1);
    else
        pd(:,1)=Pr(:,1);
        for j = 2 : n
            pd(:,j)=Pr(:,j)+pd(:,j-1);
        end
        C=rand;
        D=pd-C;
        E=find(D>0,1);
        y(i+1)=E;

        T(i+1) = T(i) + minutes(1);
    end

end
```

```
i=PMatrix_Index;
zz=Loc(i, :);
    zz(:,1) = [];
    mat_max_index{i,:}=find(zz == max(zz));
    [~,ncols(i)] = cellfun(@size,mat_max_index(i,1));
    if ncols(i) > 1
        if i==1
            L(i)=5;
        else
            L(i)=L(i-1);
        end
    else
        L(i)=mat_max_index{i};
    end


output=[L;y];
output=output.';

end
```

Appendix B – MATLAB Code of Cross-validation Process

```
function s = IFM_Val7_DiffRes_Shide_PredictionVsTestData(~)
clc
H=10;
qq=1;
s= zeros(1440,2*H);
for hh = 1:10
    month= '04Months';
    cvnumber=hh;
    res=5;
    x = {month, '-Shide-S-P-AllInstances-', 'Wed'};
    x1 = {'Shide-S-P-', cvnumber, '-Wed-TestData'};

    d=x{3};
    name     = sprintf('%s%s', x{1}, x{2}, x{3});
    name1     = sprintf('%s%d%s%', x1{1}, x1{2}, x1{3});

    full_filename = fullfile(d,name);
    full_filename1 = fullfile(d,sprintf('%d', cvnumber), name1);

    P = xlsread(full_filename, 'PWholeDay-WholeYear');

    Loc = xlsread(full_filename1, 'SWholeDay-WholeYear');

    format long g

    ST=P(1,2);
    ET=P(end,2);

    ST=datetime(ST,'ConvertFrom','datenum');
    ST = datetime(ST,'Format','HH:mm');
    ST = dateshift(ST, 'start', 'minute', 'nearest');
    NearestHrST = dateshift(ST, 'start', 'hour', 'nearest');
    MinST = minute(ST);
    HrST=hour(NearestHrST);
    if MinST >= 30

    else
        HrST=HrST+1;
    end

    ET=datetime(ET,'ConvertFrom','datenum');
    ET = datetime(ET,'Format','HH:mm');
    ET = dateshift(ET, 'start', 'minute', 'nearest');
    MinET = minute(ET);
    HrET=hour(ET);

        HrDiff=HrET-HrST;
        MinDiff=MinET+(60-MinST);

        PT=(HrDiff*60)+MinDiff;
        PMatrix_Index=(PT)+1;
```

```
T(1)=ST;
Time = zeros(1, PMatrix_Index);
y = zeros(1, PMatrix_Index);

ncols= zeros(1,PMatrix_Index);
L=zeros(1,PMatrix_Index);
mat_max_index=cell(PMatrix_Index,1);
i=1;
Time(i) = rem(datenum(T(1)),1);
    zz=Loc(i, :);
    zz(:,1) = [];
        mat_max_index{i,:}=find(zz == max(zz));
        [~,ncols(i)] = cellfun(@size,mat_max_index(i,1));
            if ncols(i) > 1
                L(i)=5;
            else
                L(i)=mat_max_index{i};
            end
y (1)=L(i);
    [a,~]=find(abs(P(:,2)-Time(i))<0.00001,5);

    PR=P(a, :);
    A=PR(L(i),:);
    kk=find(A(1,:)>=0,5,'last');
    Pr=A(kk);

    [~, n] = size(Pr);
    if Pr == zeros(1,n)
        Predicted_L=1;
        A=PR(Predicted_L,:);
        kk=find(A(1,:)>=0,5,'last');
        Pr=A(kk);
          if Pr == zeros(1,n)
             while true
                 Predicted_L = Predicted_L+1;
                 A=PR(Predicted_L,:);
                 kk=find(A(1,:)>=0,5,'last');
                 Pr=A(kk);
                    if any(Pr > zeros(1,n))
                        break;
                    end
              end
          end
        pd(:,1)=Pr(:,1);
        for h = 2 : n
            pd(:,h)=Pr(:,h)+pd(:,h-1);
        end
        C=rand;
        D=pd-C;
        E=find(D>0,1);
        y(i+1)=E;
        T(2) = T(1) + minutes(1);
    else

        pd(:,1)=Pr(:,1);
```

```
        for h = 2 : n
            pd(:,h)=Pr(:,h)+pd(:,h-1);
        end
        C=rand;
        D=pd-C;
        E=find(D>0,1);
        y(i+1)=E;

        T(2) = T(1) + minutes(1);
    end

j=2;
q=res-2;
while true
for i = j : j+q
    Time(i) = rem(datenum(T(i)),1);
    zz=Loc(i, :);
    zz(:,1) = [];
    mat_max_index{i,:}=find(zz == max(zz));
    [~,ncols(i)] = cellfun(@size,mat_max_index(i,1));
    if ncols(i) > 1
        if i==1
            L(i)=5;
        else
            L(i)=L(i-1);
        end
    else
        L(i)=mat_max_index{i};
    end
    [a,~]=find(abs(P(:,2)-Time(i))<0.00001,5);
    PR=P(a, :);
    A=PR(y(i),:);
    kk=find(A(1,:)>=0,5,'last');
    Pr=A(kk);
    [~, n] = size(Pr);

    if Pr == zeros(1,n)
        Predicted_L=1;
        A=PR(Predicted_L,:);
        kk=find(A(1,:)>=0,5,'last');
        Pr=A(kk);
          if Pr == zeros(1,n)
             while true
                 Predicted_L = Predicted_L+1;
                 A=PR(Predicted_L,:);
                 kk=find(A(1,:)>=0,5,'last');
                 Pr=A(kk);
                    if any(Pr > zeros(1,n))
                        break;
                    end
               end
          end
        pd(:,1)=Pr(:,1);
        for h = 2 : n
            pd(:,h)=Pr(:,h)+pd(:,h-1);
        end
```

```
            C=rand;
            D=pd-C;
            E=find(D>0,1);
            y(i+1)=E;
            T(i+1) = T(i) + minutes(1);
        else
            pd(:,1)=Pr(:,1);
            for h = 2 : n
                pd(:,h)=Pr(:,h)+pd(:,h-1);
            end
            C=rand;
            D=pd-C;
            E=find(D>0,1);
            y(i+1)=E;
            T(i+1) = T(i) + minutes(1);

        end
    end

    if i < PMatrix_Index-1
        if i == ((floor(PMatrix_Index/res))*res)
        zz=Loc(i+1, :);
        zz(:,1) = [];
        mat_max_index{i+1,:}=find(zz == max(zz));
        [~,ncols(i+1)] = cellfun(@size,mat_max_index(i+1,1));
        if ncols(i+1) > 1
            if i==1
                L(i+1)=5;
            else
                L(i+1)=L(i);
            end
        else
            L(i+1)=mat_max_index{i+1};
        end
        y(i+1)=L(i+1);
        q=res-1;
        j=PMatrix_Index-q;
        else
        zz=Loc(i+1, :);
        zz(:,1) = [];
        mat_max_index{i+1,:}=find(zz == max(zz));
        [~,ncols(i+1)] = cellfun(@size,mat_max_index(i+1,1));
        if ncols(i+1) > 1
            if i==1
                L(i+1)=5;
            else
                L(i+1)=L(i);
            end
        else
            L(i+1)=mat_max_index{i+1};
        end
        y(i+1)=L(i+1);
        j=i+1;
        q=res-1;
        end
    else
```

```
        break;
    end
    end
    y(end)=[];

    output=[L;y];
    output=output.';

    s(:,qq:qq+1)=output;
    qq=qq+2;
end

end
```