Summer 7-10-2021

Dynamic HVAC Energy Management Using Commercial Building Occupancy Metrics & Neural Networks

Krishna Chaitanya Jagadeesh Simma

Krishna Chaitanya Jagadeesh Simma

Candidate

Civil, Construction & Environmental Engineering

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Susan M Bogus , Chairperson

Fernando Moreu

Thomas P Caudell

Amy Ballard

DYNAMIC HVAC ENERGY MANAGEMENT USING COMMERCIAL BUILDING OCCUPANCY METRICS & NEURAL NETWORKS

by

KRISHNA CHAITANYA JAGADEESH SIMMA

B.Tech., Civil Engineering, JNTU, 2010
M.S., Civil Engineering, University of New Mexico, 2016

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Engineering

The University of New Mexico
Albuquerque, New Mexico

July 2021

Dedication

To my mother, father, sister, and my better-half Radha Swaminathan.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Dr. Susan Halter Bogus who supported me through some of my toughest times and mentored me towards my goals. I would also like to thank Prof. Thomas P. Caudell who worked patiently with me over the past two years to help me make invaluable contributions through my research. I would also like to thank my committee members Dr. Fernando Moreu, and Dr. Amy Ballard for extending their support in this journey. I would also thank Dr. Andrea Mammoli, and Richard Burnett for their significant contributions without whom this work would not have been possible. I would also like to acknowledge UNM IT department for sharing the data without which this research could not have been completed. I would also like to thank my Construction Management Lab members for their support and specifically Dr. Claudia Garrido Martins for the numerous stimulating conversations. Finally, I would like to thank Mr. Micheal Gonzalez for his support during this entire journey.

DYNAMIC HVAC ENERGY MANAGEMENT USING BUILDING OCCUPANCY METRICS AND NEURAL NETWORKS

by

Krishna Chaitanya Jagadeesh Simma

B.Tech., Civil Engineering, JNTU Kakinada, 2010

M.S., Civil Engineering, University of New Mexico, 2016

PhD., Engineering, University of New Mexico, 2021

ABSTRACT

With the rise of technology use in buildings, it is now possible to collect data that can be used to improve building energy consumption. One factor that has significant impact on building energy consumption is occupancy. Recent studies have shown promising results in obtaining occupancy information from existing infrastructure such as WiFi router networks. However, these existing frameworks require additional investments through software upgrades, added infrastructure, computational resources, and may raise occupant privacy concerns. Additionally, with occupant thermal comfort statistics being lower than ASHAREA specified standards, a novel approach for indoor climate control is needed. To address the limitations in existing frameworks and the lower occupant thermal comfort statistics, this study proposes a framework to estimate occupancy from existing WiFi network data with minimal computational efforts and reduced privacy concerns. The structure of the occupancy data was studied to learn patterns within the data. The learned patterns are then used to make short term occupancy profile predictions using neural networks. The WiFi measured occupancy data showed a correlation up to 0.96 implying that WiFi client-count can be a reliable source of occupant count. HVAC energy consumption values were estimated using simulation models developed in EnergyPlus for the lecture hall used for data collection. The HVAC energy consumption for different occupancy-based schedules was

estimated and compared against a fixed schedule (typical 6am-6pm operation assuming full occupancy), and a registered schedule (using total number registered occupants per each lecture scheduled in the lecture hall). The energy consumption results suggest that occupant demand-driven HVAC operation can results in energy savings of 50% compared to the fixed schedules.

The results from the occupancy pattern study revealed that by dividing the 24-hour occupancy profiles into smaller segments bound by external schedules, significant patterns can be learned. In this study, the external schedules were defined by the length of lectures scheduled for the Fall 2019 semester. Additionally, each time-segment will have at least one frequently occurring pattern labeled as Most Likely Template (MLT) and a pattern with zero-occupancy profiles labeled as Holiday Template (HT). The MLT patterns identified per time-segment can serve as expected occupancy for that time-segment. However, the different patterns learned for each time-segment suggest that using MLT profiles for expected occupancy may not always be correct. Therefore, a prediction mechanism was devised to avoid incorrect profile assignments for a time-segment. Using the patterns learned and neural networks (Laterally Primed Adaptive Resonance Theory, LAPART), short term occupancy profile predictions were made. The results from the prediction analysis imply that reliable occupancy prediction is possible if there are intrinsic or extrinsic variables with significant correlation to occupancy. The forecast of occupant count can be used for pre-conditioning the space personalized to the number of occupants in the space (i.e., establishing thermal setpoints as a function of occupant count) which can increase occupants' thermal comfort and reduce unnecessary HVAC demand.

Keywords

WiFi; Commercial buildings, Occupancy; Prediction; HVAC; Energy Efficiency; Demand-Driven; Neural Networks; Fuzzy ART; LAPART.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Chapter-01: Introduction

1. Introduction

Commercial and residential building energy consumption accounts for 45% of the total energy consumption in the U.S as illustrated in **Figure 1** (EIA 2019). The latest Energy Information Administration (EIA) projections suggest that residential building energy consumption is down and will continue the downward trend for the next two decades (EIA 2019). On the contrary, commercial building energy consumption is expected to grow for the next three years at a rate of 0.5% per year. The aging commercial building stock in the U.S. (CBECS, 2012) provides an opportunity to develop novel frameworks to reduce commercial building energy consumption.

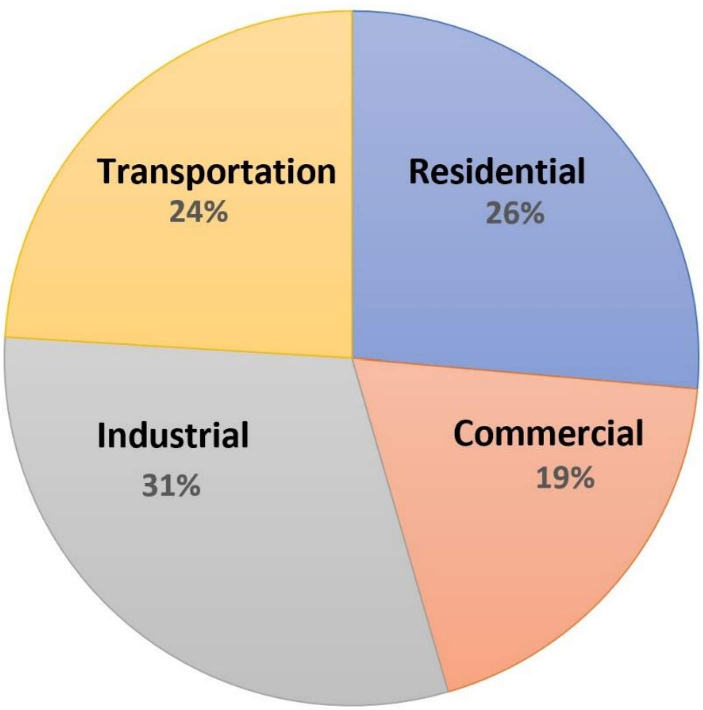


Figure 1: Total Energy Consumption distribution in the U.S. (EIA, 2019)

Unlike residential buildings that are typically occupied by fewer 'long-term' occupants, commercial building occupants are higher in number and stochastic in nature. The complexity in commercial building occupancy has led to significant deviations in actual energy consumption compared to modeled energy consumption (Azar and Menassa 2012a). To address these deviations, numerous frameworks were proposed to collect, analyze, and model occupancy in commercial buildings. As the technology integration into commercial buildings grew, novel frameworks were proposed to detect, estimate, and track occupants and occupant movement within the buildings. Collecting occupancy data often required dedicated infrastructure which can be expensive for building owners. This led to a set of studies that proposed occupancy data collection from existing infrastructure. Based on the infrastructure used, existing frameworks can be classified into two sets: 1) frameworks that use dedicated infrastructure and 2) frameworks that use existing infrastructure. Studies such as Fleuret et al (2008), N. Li et al (2012), Pan et al (2014), and Labeodan et al (2015) collected data from chair sensors, vibratory sensors, cameras, and RFID tags to obtain occupant information. Existing infrastructure frameworks used humidity sensors, thermostats, bluetooth, and commodity WiFi to obtain occupancy information (Ekwevugbe et al. 2013; M. Wang et al. 2014; Ekwevugbe et al. 2017). Both the listed frameworks collected occupancy data at different levels with specific objectives. Three different levels of occupancy were identified from the existing literature: 1) occupancy detection, 2) occupancy estimation, and 3) occupancy tracking.

The costs associated with dedicated infrastructure for occupancy data collection led researchers to favor methods that used existing infrastructure. Commodity WiFi gained popularity in the past few years due to its availability in most commercial buildings. The current WiFi based frameworks predominantly used two metrics to obtain occupancy data: 1) Received

Signal Strength (RSS) and 2) Channel Slate Information (CSI). These WiFi based metrics were used in frameworks such as WinOSS (Zou et al. 2017), WiFree (Zou et al. 2018), FreeDetector(Zou et al. 2017) , FreeCount (Zou et al. 2018), and WiFi Pineapple (Çiftler et al. 2018). While these frameworks offer valuable evidence to the utility of commodity WiFi as a source of occupancy data, they often require additional infrastructure, upgrades, and cumbersome computational requirements. Additionally, the data collected consists of unique occupant identifiers that raises privacy concerns. Furthermore, these frameworks were implemented in highly controlled office spaces with limited occupancy and may not reflect the occupant dynamics of large commercial buildings such as airports, libraries, institutional buildings, and gymnasiums. These shortcomings may hinder large scale implementation of these frameworks and the realization of the full energy saving potential in large commercial buildings with dynamic occupancy.

In commercial buildings, a substantial amount of energy is spent for Heating, Ventilation and Air Conditioning (HVAC) systems. Typically, commercial building HVAC systems are operated on static/fixed schedules (e.g., 6am – 6pm). Often these schedules do not include the real-time occupant loads and literature suggests that operating HVAC systems based on occupancy schedules can result in significant energy savings (Z. Yang and Becerik-Gerber 2014). Additionally, obtaining accurate and high-resolution occupancy information ahead of time allows for better management of building energy by avoiding wasteful HVAC demand (W. Wang et al. 2018). To this extent, existing frameworks collected data from office spaces (e.g., office spaces in commercial and institutional buildings) to make occupancy predictions (Peng et al. 2018; W. Wang et al. 2018; W. Wang et al 2018). The occupants in these office spaces were assigned a unique identifier (e.g., MAC address) and using the identifiers the occupant’s state (i.e., ‘in’ or ‘out’) was predicted and consequently the total occupancy of the space. However, this approach is limited to commercial spaces which are frequently occupied by the same occupants (aka: ‘long-term’ occupants). Occupants of large commercial buildings such as airports, shopping malls, and gymnasiums cannot be classified as ‘long-term’ occupants and assigning identifiers to occupants is not only infeasible but may also raise occupant privacy concerns. Furthermore, only 39% of occupants in commercial office buildings in North America reported to be satisfied with building temperatures which is significantly lower than ASHARE standard (i.e., at least 80%) (D. Li et al. 2017).

The heat dissipation of occupants inside buildings contribute towards the cooling load in the cooling dominated months and (Q. Wang et al. 2016; W. Wang et al. 2017). This occupant heat dissipation effect has profound impact on heating loads in heating dominated months (Q. Wang et al. 2016). Additionally, the metabolic rates of occupants vary depending on the type of commercial building. The metabolic rates of occupants in gymnasiums, stores, and terminal buildings are reported to be higher at 1.6met (metabolic equivalent unit) followed by schools at 1.2met compared to other types of commercial buildings (Ahmed et al. 2017). The higher metabolic rates induce higher heat loss of occupants due to their homotherm nature (Van Marken Lichtenbelt and Kingma, 2013). The metabolic rates and the amount of heat dissipated by occupants in both summer and winter months are identical (Ahmed et al., 2017), highlighting that occupants’ thermal interaction in all types of weathers remain constant. Furthermore, human body radiates different levels of heat throughout the day, implying that the occupants’ thermal interactions within buildings are dynamic throughout the day (Mege, 2021). Moreover, dynamic variation of indoor spatial temperature that accounts of thermo-physiological parameters such as occupants’ metabolic rate and sweat production can positively impact their comfort and health

(Van Marken Lichtenbelt and Kingma, 2013). In summary, a constant heating or cooling setpoint temperature may not improve occupants' thermal comfort inside buildings. New setpoints that are a function of number of occupants in a space may contribute towards improving occupant comfort inside buildings. Additionally, knowing the occupant count for a space ahead in time can allow for pre-conditioning the space to that specific number of people can result in increased occupant comfort and avoid HVAC wasteful demand. With emphasis on the need for such HVAC strategies (Klein et al., 2012; Liao and Barooah, 2010), an occupancy prediction approach that addresses the identified limitations is essential.

This research presents a novel approach for accurate occupancy prediction using data collected from existing infrastructure and reflects the dynamics of a large commercial space that does not utilize occupant identifiers. To this extent, this study includes a feasibility test where identifier-free occupancy data were collected for a relative dynamic environment (i.e., a large lecture hall) using the university WiFi network. These occupancy data were analyzed to learn patterns of repetition using neural networks and the patterns learned were used to make short term occupancy predictions. HVAC energy consumption was estimated using EnergyPlus models at different stages to evaluate the potential energy savings and to validate the occupancy data.

2. Research Questions

The goal of this research is to use WiFi networks to obtain reliable occupancy data and detect patterns within the data to make short term predictions that can be used to reduce HVAC energy consumption and increase occupant comfort levels. This is achieved by answering the following questions:

1. How accurately does the WiFi ‘client-count’ data represent the actual occupant in a relative dynamic environment?
2. What amount of HVAC energy savings can be achieved using real-time WiFi occupancy data from a relative dynamic environment?
3. How can patterns of repetition be recognized in the WiFi occupancy data?
4. How should the learned patterns be used to make short term occupancy predictions?

These questions are answered by the achieving the following objectives

1. Estimate the accuracy of WiFi client count to establish it as a reliable source of occupant data for this study.
2. Estimate the potential energy savings of WiFi-based occupancy schedules over static occupancy schedules.
3. Analyze occupancy data to recognize patterns using neural networks.
4. Predict occupancy with variables that are significantly correlated using neural networks.

3. Methodology

The methodology for this study is divided into four parts, one per objective: 1) WiFi occupancy estimation: An occupancy counter was setup to establish the ground truth for occupancy. The WiFi data was processed and validated for accuracy, 2) HVAC energy estimation: Different occupancy schedules were created for HVAC operation and the associated energy consumption was estimated using EnergyPlus models, 3) Occupancy pattern recognition: Patterns in WiFi occupancy were learned using a Fuzzy ART neural network, and 4) Occupancy profile prediction: Using the patterns learned, occupancy profile predictions were made with LAPART neural networks. **Figure 2** graphically illustrates the methodology.

This dissertation is divided into six chapters starting with Chapter 1 that introduces the area of study and ends with Chapter 6 that provides an overall conclusion, contribution to the existing body of knowledge, and future directions. Chapters 2, 3, 4, and 5 form the main body of this dissertation and each chapter corresponds to an objective listed in the methodology**.** Chapter 2 has been presented and published in the proceedings of *Canadian Society for Civil Engineering Annual conference, 2019.* This chapter establishes the WiFi data as an accurate source of occupancy for large commercial buildings. Chapter 3 has been published in *Procedia Computer Science, 2019 (*https://doi.org/10.1016/j.procs.2019.08.069). This chapter establishes that occupancy-based schedules result in significant HVAC energy savings compared to static/fixed schedule. Chapter 4 was presented at the *Canadian Society for Civil Engineering Annual Conference*, 2021 and will be published in *Springer Nature*, 2021. This chapter relates to the investigation of pattern recognition in the occupancy data collected from the WiFi networks and Fuzzy ART neural networks. Chapter 5 presents the results of the feasibility test that offers evidence to the possibility of short-term occupancy prediction using intrinsic variables that are highly correlated to the occupancy profiles. The manuscript resulting from this chapter will be submitted to the journal *Energy and Buildings*. Thus, the fundamental knowledge gained from this research suggests that WiFi device count can accurately represent occupancy in a relative dynamic environment. Significant occupancy patterns can be learned from the WiFi data by fragmenting the data into smaller time-segments bound by external schedules. Additionally, the prediction feasibility test results lend evidence to the possibility of predicting occupancy profiles for a time-segment using variables that are significantly correlated to the occupancy of a time-segment. Furthermore, the patterns learned, and the subsequent predictions can aid in realizing substantial HVAC energy savings and can also be used towards improving occupant comfort levels.

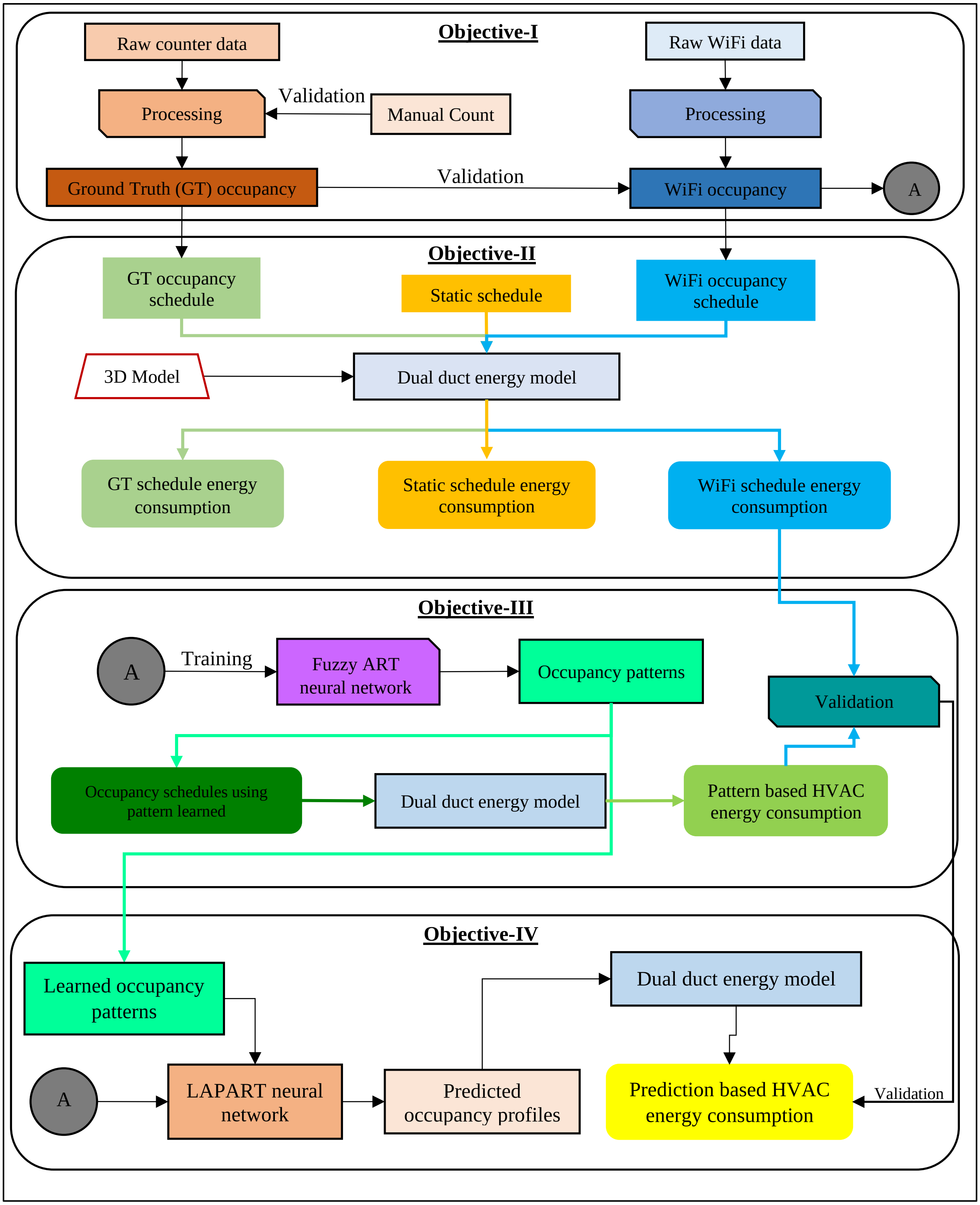


Figure 2: Overall Methodology

References

Ahmed, K., Kurnitski, J., & Olesen, B. (2017). Data for occupancy internal heat gain calculation in main building categories. *Data in Brief*, *15*, 1030–1034. https://doi.org/10.1016/j.dib.2017.10.036

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. https://doi.org/10.1016/j.enbuild.2012.10.002

Çiftler, B. S., Dikmese, S., Güvenç, I., Akkaya, K., & Kadri, A. (2018). Occupancy Counting With Burst and Intermittent Signals in Smart Buildings. *IEEE Internet of Things Journal*, *5*(2), 724–735. https://doi.org/10.1109/JIOT.2017.2756689

EIA, & Hojjati, B. (2019). Global energy consumption driven by more electricity in residential, commercial buildings. *U.S. Energy Information Association*, *2019*, 2019–2020. Retrieved from https://www.eia.gov/todayinenergy/detail.php?id=41753

Ekwevugbe, T., Brown, N., Pakka, V., & Fan, D. (2013). Real-time building occupancy sensing using neural-network based sensor network. In *2013 7th IEEE International Conference on Digital Ecosystems and Technologies (DEST)* (pp. 114–119).

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

Fleuret, F., Berclaz, J., Lengagne, R., & Fua, P. (2008). Multicamera people tracking with a probabilistic occupancy map. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(2), 267–282.

Klein, L., Kwak, J. Y., Kavulya, G., Jazizadeh, F., Becerik-Gerber, B., Varakantham, P., & Tambe, M. (2012). Coordinating occupant behavior for building energy and comfort management using multi-agent systems. *Automation in Construction*, *22*, 525–536. https://doi.org/10.1016/j.autcon.2011.11.012

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*, 303–314. https://doi.org/10.1016/j.enbuild.2015.02.028

Li, D., Menassa, C. C., & Kamat, V. R. (2017). Personalized human comfort in indoor building environments under diverse conditioning modes. *Building and Environment*, *126*(September), 304–317. https://doi.org/10.1016/j.buildenv.2017.10.004

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*. https://doi.org/10.1016/j.autcon.2012.02.013

Liao, C., & Barooah, P. (2010). An Integrated Approach to Occupancy Modeling and Estimation in Commercial Buildings, 3130–3135.

Mege, J. (2021). Variations , and the Immune System.

Pan, S., Bonde, A., Jing, J., Zhang, L., Zhang, P., & Noh, H. Y. (2014). BOES: Building Occupancy Estimation System using sparse ambient vibration monitoring. *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, 90611O. https://doi.org/10.1117/12.2046510

Peng, Y., Rysanek, A., Nagy, Z., & Schlüter, A. (2018). Using machine learning techniques for occupancy-prediction-based cooling control in office buildings. *Applied Energy*, *211*(November 2017), 1343–1358. https://doi.org/10.1016/j.apenergy.2017.12.002

Van Marken Lichtenbelt, W. D., & Kingma, B. R. (2013). Building and occupant energetics: A physiological hypothesis. *Architectural Science Review*, *56*(1), 48–53. https://doi.org/10.1080/00038628.2012.759377

Wang, M., Wang, X., Zhang, G., & Li, C. (2014). Occupancy detection based on spiking neural networks for green building automation systems. In *Proceeding of the 11th World Congress on Intelligent Control and Automation* (pp. 2681–2686).

Wang, Q., Augenbroe, G., Kim, J. H., & Gu, L. (2016). Meta-modeling of occupancy variables and analysis of their impact on energy outcomes of office buildings. *Applied Energy*, *174*, 166–180. https://doi.org/10.1016/j.apenergy.2016.04.062

Wang, W., Chen, J., & Hong, T. (2018). Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings. *Automation in Construction*, *94*(July), 233–243. https://doi.org/10.1016/j.autcon.2018.07.007

Wang, W., Chen, J., Hong, T., & Zhu, N. (2018). Occupancy prediction through Markov based feedback recurrent neural network (M-FRNN) algorithm with WiFi probe technology. *Building and Environment*, *138*(March), 160–170. https://doi.org/10.1016/j.buildenv.2018.04.034

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*, *207*, 305–323. https://doi.org/10.1016/j.apenergy.2017.06.060

Yang, Z., & Becerik-Gerber, B. (2014). The coupled effects of personalized occupancy profile based HVAC schedules and room reassignment on building energy use. *Energy and Buildings*, *78*, 113–122. https://doi.org/10.1016/j.enbuild.2014.04.002

Zou, H., Jiang, H., Yang, J., Xie, L., & Spanos, C. J. (2017). Non-intrusive occupancy sensing in commercial buildings. *Energy and Buildings*, *154*(September), 633–643. https://doi.org/10.1016/j.enbuild.2017.08.045

Zou, H., Zhou, Y., Yang, J., Gu, W., Xie, L., & Spanos, C. (2017). FreeDetector: Device-Free Occupancy Detection with Commodity WiFi. *2017 IEEE International Conference on Sensing, Communication and Networking, SECON Workshops 2017*. https://doi.org/10.1109/SECONW.2017.8011040

Zou, H., Zhou, Y., Yang, J., Gu, W., Xie, L., & Spanos, C. (2018). FreeCount: Device-free crowd counting with commodity WiFi. *2017 IEEE Global Communications Conference, GLOBECOM 2017 - Proceedings*, *2018-Janua*, 1–6. https://doi.org/10.1109/GLOCOM.2017.8255034

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. https://doi.org/10.1016/j.enbuild.2018.06.040

Chapter-02 – WiFi Router network-based occupancy estimation to optimize HVAC energy consumption

Krishna Chaitanya J Simma[1], Susan M Bogus[2], and Andrea Mammoli[3]

[1]PhD student, Department of Civil, Construction and Environmental Engineering, University of New Mexico, MSC01 1070, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: jagadeesh145@unm.edu

[2]Professor, Department of Civil, Construction and Environmental Engineering, University of New Mexico, MSC01 1070, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: sbogus@gmail.com

[3]Professor, Department of Mechanical Engineering, University of New Mexico, MSC01 1070, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: mammoli@unm.edu

ABSTRACT

More than half of the commercial building stock in the United States was built before 1980 prior to the increased focus on energy efficiency. In the current age of Smart and Green buildings, owners incorporating expensive sensor infrastructure to reduce building energy consumption and improve the building occupants' satisfaction, efficiency, and comfort levels. The success of these automated building systems is influenced by the ability to estimate building occupancy. Recently, researchers shifted their focus towards exploring different occupancy estimation techniques with both dedicated sensors and existing infrastructure (e.g., CO_2 sensors, Smart meters, temperature, and humidity sensors, and wi-fi networks). However, there are concerns about the cost effectiveness, computational effort, accuracy, and privacy protection for these

techniques. This study explores the usage wi-fi router data to generate the of number of IP addresses connected to the router to estimate the occupancy within a building. To this end, occupancy patterns in a thirty-year-old university building are estimated using existing wi-fi infrastructure and compared and calibrated to ground data obtained manually and from dedicated occupancy estimating sensors to evaluate the accuracy. The estimated occupancy data patterns using existing wi-fi network represent a cost-effective method of occupancy estimation with less computational processing and reduced privacy concerns, that could assist owners in the decision-making process towards investing into smart and energy efficient technologies.

1. Introduction

With the rise of technology, Smart buildings and green initiatives have grown in the past few years. In 2011, a report from the United States Energy Information Administration (EIA) reported an increase in the number of pilot studies related to smart grids. It stated that the smart meter installations in the United States would exceed 80 million by the year 2015 (SAIC, 2011). This is close to the EIAs' 2016 reported value of above 70 million smart meter installations in the residential, commercial, industrial, and transportation sectors. Although the current number is slightly behind the predicted value, it is evident that building owners are investing in smart technologies to improve efficiency and comfort. With more than half of commercial building stock in the US being over 32 years old (CBECS 2012), the potential for building owner's investment into smart technologies to optimize energy consumption and improve occupant comfort is great.

Commercial buildings consume about 19% of total energy consumption in the US (Azar and Menassa 2014) in which about 50% of energy is consumed by HVAC (heating, ventilation, and air conditioning) equipment. Energy models and predictions were often mismatched with the actual building performances in terms of their energy consumption. Often the mismatch between modelled energy consumption and actual energy consumption in commercial buildings is attributed to the occupants and occupant behavior of the buildings (Azar and Menassa 2012b). In the past decade, studies have emphasized on the impact of occupants on building energy consumption (Yang and Wang 2013, Labeodan et al. 2015, Hong et al. 2016). As the influence of occupants on building energy consumption became evident, the importance of occupancy information has become the point of interest for researchers.

Numerous occupancy detection and estimation techniques were introduced over the past few years. Studies have explored different techniques to detect, estimate and track occupants within the building. Some of the techniques include but are not limited to usage of sensor networks such as passive infrared sensors (PIR) (Dodier et al. 2006), RFID tags (Li et al. 2012), occupancy sensors and motion detectors (Duarte et al. 2013, Stoppel and Leite 2014, Mantha et al. 2015), vibration sensors (Pan et al., 2014), chair sensors (Labeodan et al., 2015), and Ultra-wideband (UWB) (Choi et al. 2018) among others.

However, dedicated sensor infrastructure can be expensive for large scale deployment in commercial buildings. To address these issues, researchers have investigated occupancy detection, estimation and tracking for multiple purposes using existing infrastructure such as smart meters (Kleiminger et al. 2013), cameras (Liu et al. 2013), and wi-fi routers (Vattapparamban et al. 2016, Zou et al. 2017, Zou et al. 2018) among others. Each of these existing infrastructure systems have different levels of detection and estimation accuracies and privacy concerns. Z.Chen et al. (2018), performed a comparative review of different occupancy sensing techniques. The article presented a summary of different types of sensors used to detect and estimate occupancy along with their limitations. Overall, from literature is it evident occupancy data can be categorized into three levels depending on the extent of information obtained: 1) detection, 2) estimation, and 3) location tracking (Zou et al. 2017).

This paper focuses only on occupancy estimation using existing infrastructure. Infrastructure such as smart meters are capable of detecting occupancy but have no capability of estimating the occupancy (D. Chen et al. 2013). Cameras have high accuracy, however they have high computational requirements and privacy concerns which would restrict their usage (Liu et al. 2013Z). Wi-fi signals are capable of detecting and estimating occupancy with partial privacy concerns of the occupants (Zou et al. 2017, Zou et al. 2018). However, from the studies on occupancy estimation using wi-fi routers/Access Points (AP's) and signals (Received Signal Strength, RSS), it is evident that the occupancy estimation requires either significant computational resources, additional software updates to the routers, or additional devices installed (Depatla et al. 2015, Vattapparamban et al. 2016, Zou et al. 2017). Table 1 summarizes some of the occupancy estimation techniques proposed in recent literature along with their computational requirements, added infrastructure, reported accuracy, and concerns.

Table 1: Summary of Wi-Fi based Occupancy Estimation Methods

Name	Infrastructure Used	Additional Resources	Accuracy reported	Concerns	Source
WinOSS	Wi-Fi	Firmware upgrades	98.85% (detection only)	Occupant identification	(Zou et al. 2017)
WiFree	Wi-Fi	Second Wi-Fi Router	92.80%	Computational requirements	(Zou et al. 2018)
Meraki	Wi-Fi	Meraki wireless APs	-	Occupant identification	(Cisco, 2013)
Wi-Fi Pineapple	Wi-Fi	Wi-Fi Sniffers	-	Occupant identification	(Vattapparamban et al. 2016)
FreeDetector	Wi-Fi	Firmware upgrades	94.0% (detection only)	Computational requirements	(Zou et al. 2017)

From the summary presented in Table 1, it is evident that the techniques implemented to detect and estimate occupancy require additional resources such as routers capable of handling specific task (e.g. Meraki routers), upgrading firmware, and wi-fi sniffers (e.g. wi-fi pineapple) among others, identifies occupants through unique identifiers (e.g. MAC addresses), or limited to occupancy detection only. The added infrastructure, and firmware upgrades may increase the cost of gathering occupant data for commercial buildings. Similarly, identifying and tracking individuals may raise privacy concerns when implemented in university buildings or other public buildings. In this context, this paper asks a question: *Can Wi-fi Routers serve as a cost-effective, reliable, and accurate source of occupancy estimates that reduces computational requirements and privacy concerns?*

2. Methodology

To address the question asked, this study proposes the methodology presented in Figure 3 and consisting of three steps: 1) Establish ground truth, 2) Data acquisition, and 3) Data processing, and accuracy.

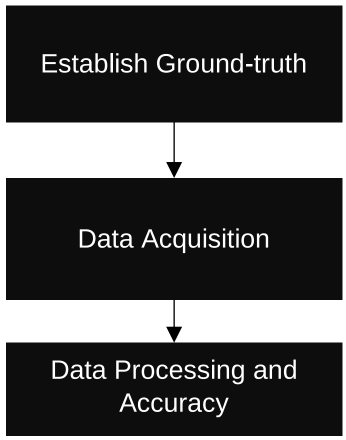


Figure 3: Methodology (Objective-I)

The methodology is used to estimate occupancy of a large lecture hall inside a thirty-year old Mechanical Engineering Building at University of New Mexico that is equipped with a campus wide wi-fi network. To estimate the occupancy, it is assumed that when students spend time within the university building, they connect to the university wi-fi network for their needs. The router infrastructure covers the entire building which facilitates the detection and estimation of occupants within the areas of wi-fi coverage. The lecture hall in question was preinstalled with three wi-fi routers spread across the entire room.

2.1. Step1: Establish Ground Truth

To establish ground truth, the lecture hall in the Mechanical Engineering building shown in Figure 4(a) was selected as it is one of the classrooms regularly used during the semester. The lecture hall is capable of seating over one hundred students at a time. It has two entrances one on the north end and one on the south end. On average five different classes take place on a regular weekday. To obtain an actual count during a normal class, a people-counting sensor (EBTRON: CENCUS-C100) shown in Figure 4(b) was installed that uses the thermal signature of occupants to estimate the occupant count as they walk through the door. Each entrance was installed with a single C100 as shown in Figure 4(c). When an individual enters through the door, the C100 sensor is activated, and it is directionally sensitive. It consists of two infrared sensors that detects the thermal signature of the occupant and increases the count when an individual enters and decreases when the individual exits based on the order of activation (e.g., if 1 to 2 is entry, 2 to 1 is exit).

(a)

(c)

(b)

Figure 4: (a)Lecture Hall under study, (b) EBTRON: CENCUS-C100 (c) Installed sensor on door frame.

The installed sensors were calibrated and tested for over 3 months during regular semester weekdays. The sensor logs the occupancy count every time an occupant walks through the door to attend a class and sends the data to the server located at the Physical Plant Department on the university campus. The occupancy count data is made available for download from the server as a Comma Separated Value (CSV) file. The raw data consists of the occupant count for the entire day with timestamps. This data is then validated against manual counts to estimate the accuracy and establish the sensor count as the ground truth.

2.2.Step2: Data Acquisition from Routers

The lecture hall is equipped with three wi-fi routers to facilitate wi-fi coverage for the entire hall. Students often connect to the wi-fi network during classes and this data is logged and sent to the network servers held at the Information Technology (IT) department for the university campus. This data consists of the number of clients (i.e., number of Media Access Control (MAC) addresses) connected to the network at a given time throughout the day. Such data can be obtained for any building equipped with a wireless network managed by a central network server. For this investigation, the IT department was asked to share the data with number of clients connected through the wi-fi routers inside the lecture hall throughout the day. The number of connections at a given time should approximately match the total occupancy of the lecture hall. The IT department was asked to filter any information that could identify an occupant to eliminate privacy concerns.

2.3.Step3: Data Processing and Accuracy

The client list is shared on daily basis as a CSV file containing the data from the previous day. This data requires minimal processing to estimate occupancy of the room as the count of total number of clients at a given time is considered as the total occupancy. This client count data is then compared with the ground truth (data obtained by the sensors (C100) installed for the lecture hall) to find the correlation between the two estimates and measure the accuracy.

3.Results and Discussion

The installed sensors were connected to the university's Delta Control systems network (Facilities management system) to allow viewing the data logging in real time as shown in Figure 5. The sensors were calibrated and tested during regular semester classes and special seminar talks where the total attendance was obtained via manual count. A data point is logged every time a student walks through the door. The count increases as students walk in through the door and decreases as students walk out. No specific instructions were given to the students on how to enter or exit the room. The student's behavior was unaltered throughout the period of calibration and testing. The logged data provides fine grained occupancy information in real time as students walk in and walk out of the lecture hall. The data is then compared to the manual count over multiple days and the sensor achieved 97.7% accuracy in estimating the occupancy count. Therefore, the sensor count is used going forward as representative of the ground truth.

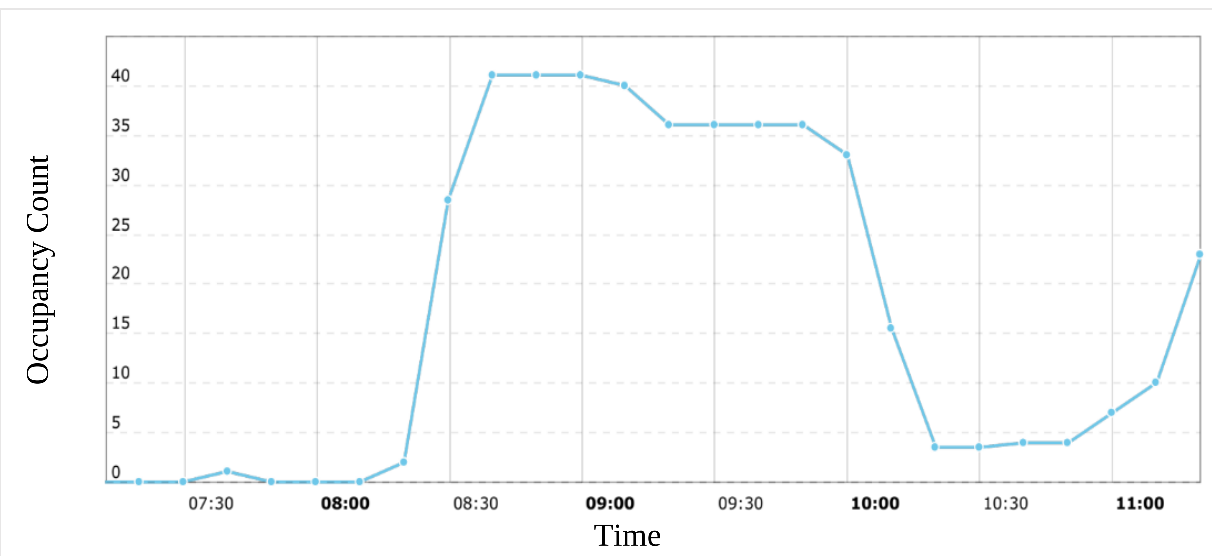
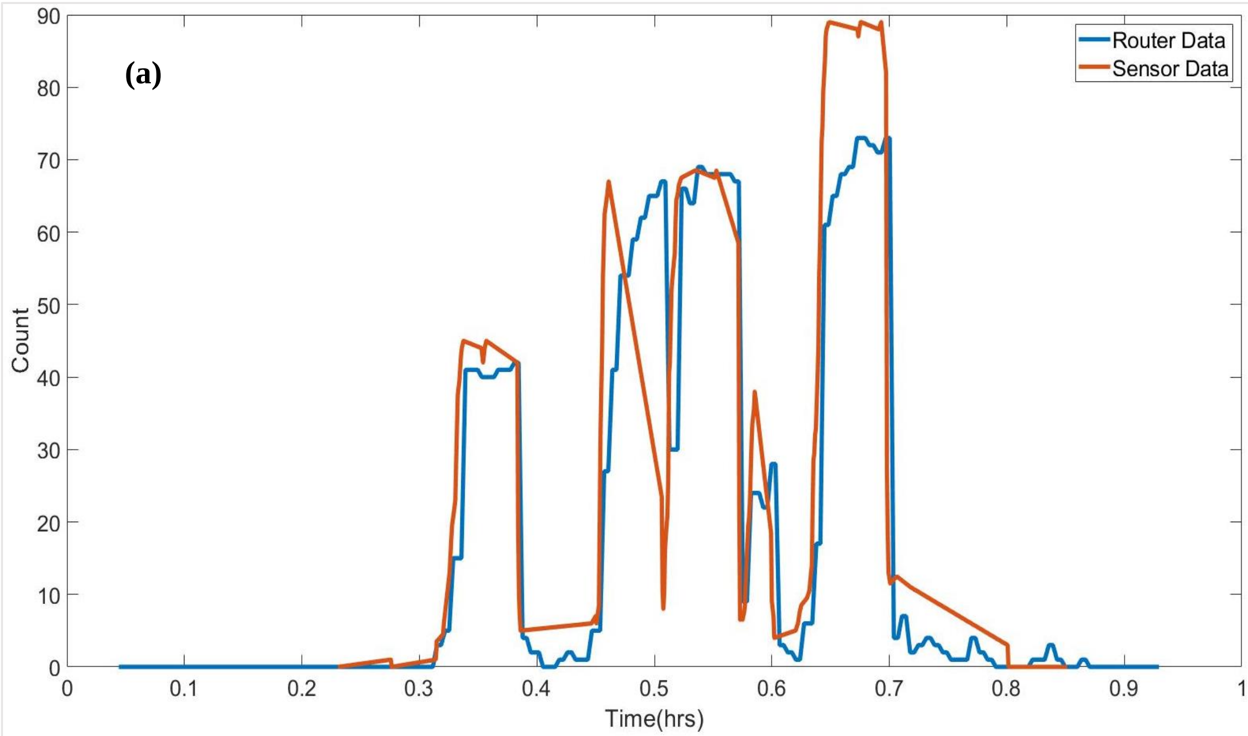
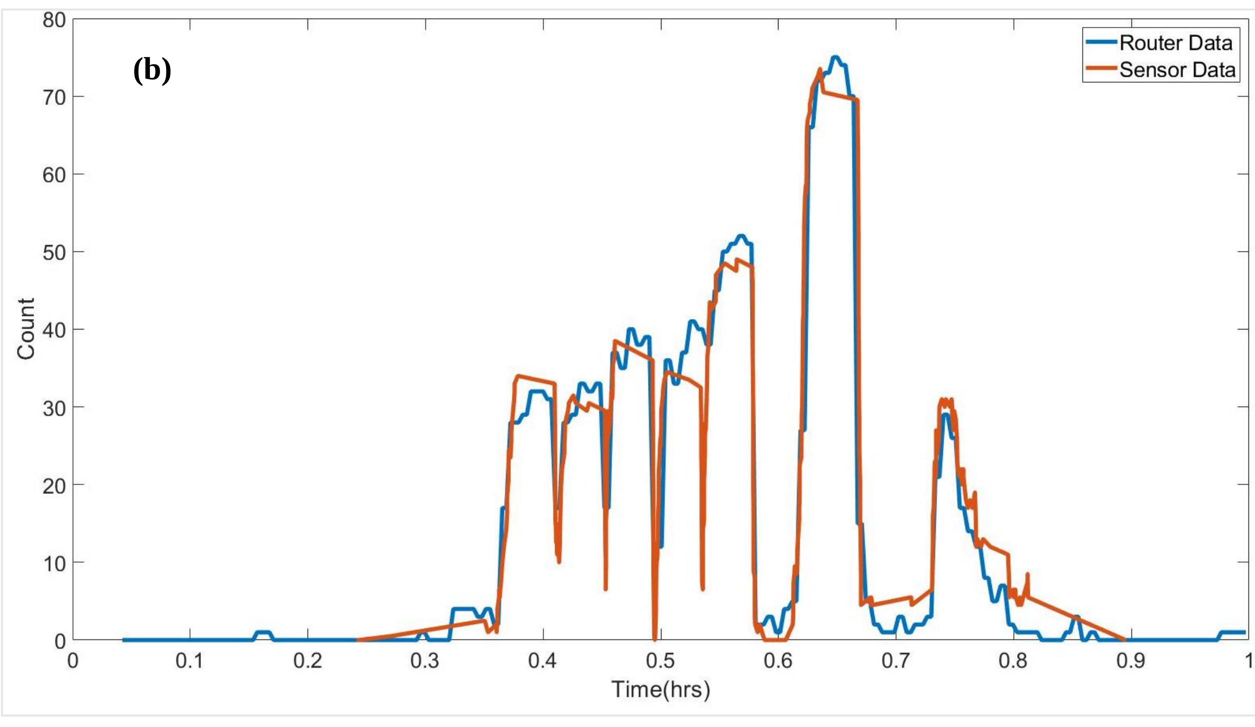


Figure 5: Data logged by the Sensor in Real Time (June 19, 2018)

The wi-fi routers in the lecture hall allow students to connect to the campus network through one of the routers and the information of the individual is logged on the university IT department's network servers. The servers log the total number of clients connected to the campus wi-fi every five minutes throughout the day. The total number of unique clients connected to the three routers that serve the lecture hall were isolated from the rest of the database with a timestamp. This information was shared via CSV file from Jan 22, 2019, to Feb 21, 2019. All the information such as MAC or IP address of the users that can identify an individual was filtered out by the IT department to protect the identity of the occupants.

The total count versus time from the two data sets are plotted alongside each other using simple MATLAB script as shown in Figure 6 from (a) to (d) representing the data from Jan 22, 2019, to Jan 25, 2019, respectively. The timesteps at which the data logged by the C100 sensor is different from that of the wi-fi routers. To form a correlation between the two data sources, the occupancy values need to be obtained for the same timesteps from each source. Using the "griddedInterpolant" function in MATLAB, occupant count and client count were interpolated for the same timesteps. The extracted values provided the occupant count ($x1$) from the C100 sensor and client count ($x2$) from the wi-fi router. As time (y) is common for both $x1$ and $x2$, these values are plotted against each other to estimate the correlation. The correlation plot with linear regression line is shown in Figure 7 and Figure 8 for days Jan 22, 2019, to Jan 25, 2019.


(a)
Router Data
Sensor Data
Count
Time(hrs)
0
10
20
30
40
50
60
70
80
90
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


(b)
Router Data
Sensor Data
Count
Time(hrs)
0
10
20
30
40
50
60
70
80
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

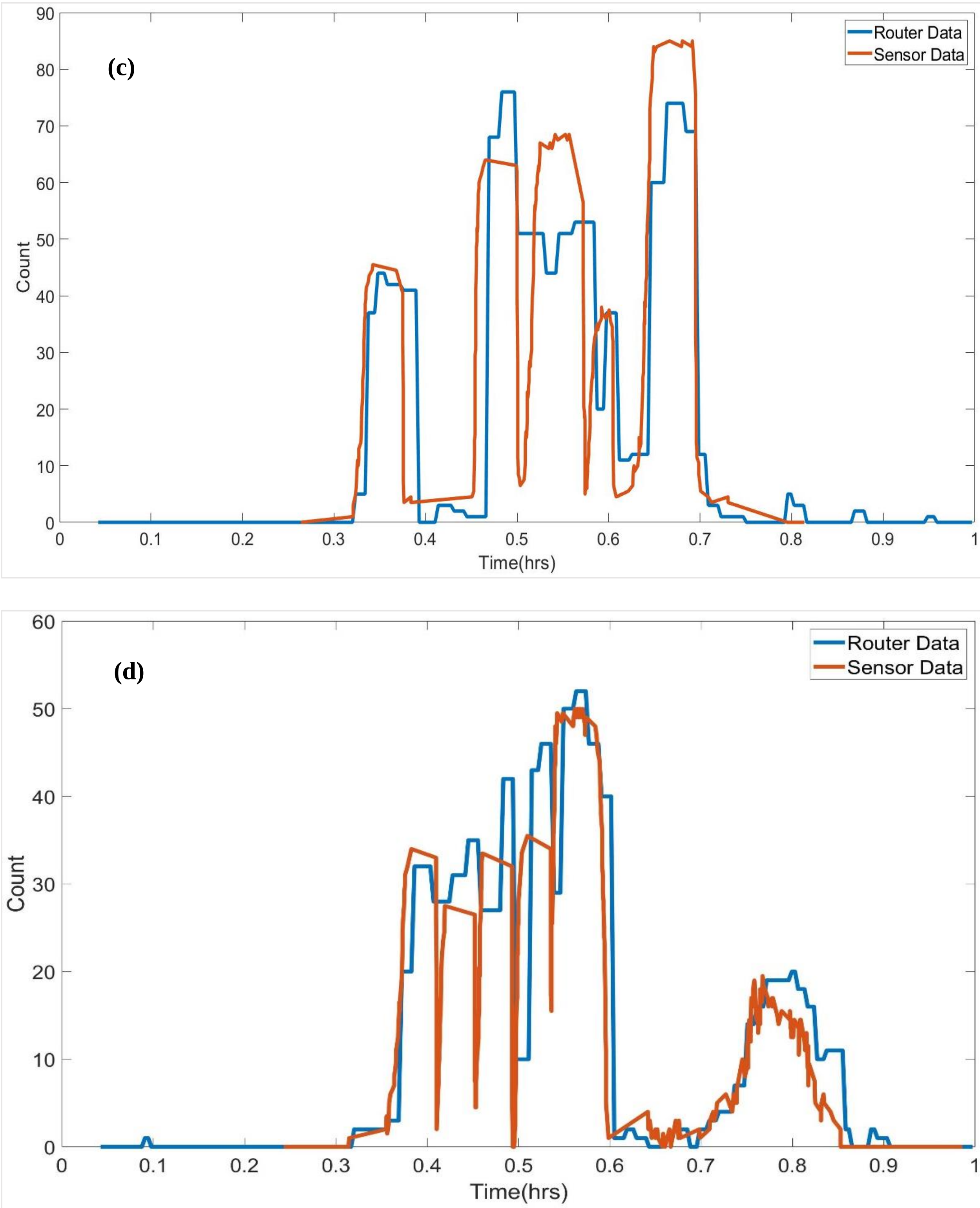


Figure 6: Sensor (C100) Occupancy count and Wi-fi Router Count vs time (a) Jan 22, 2019, (b) Jan 23, 2019, (c) Jan 24, 2019, (d) Jan 25, 2019

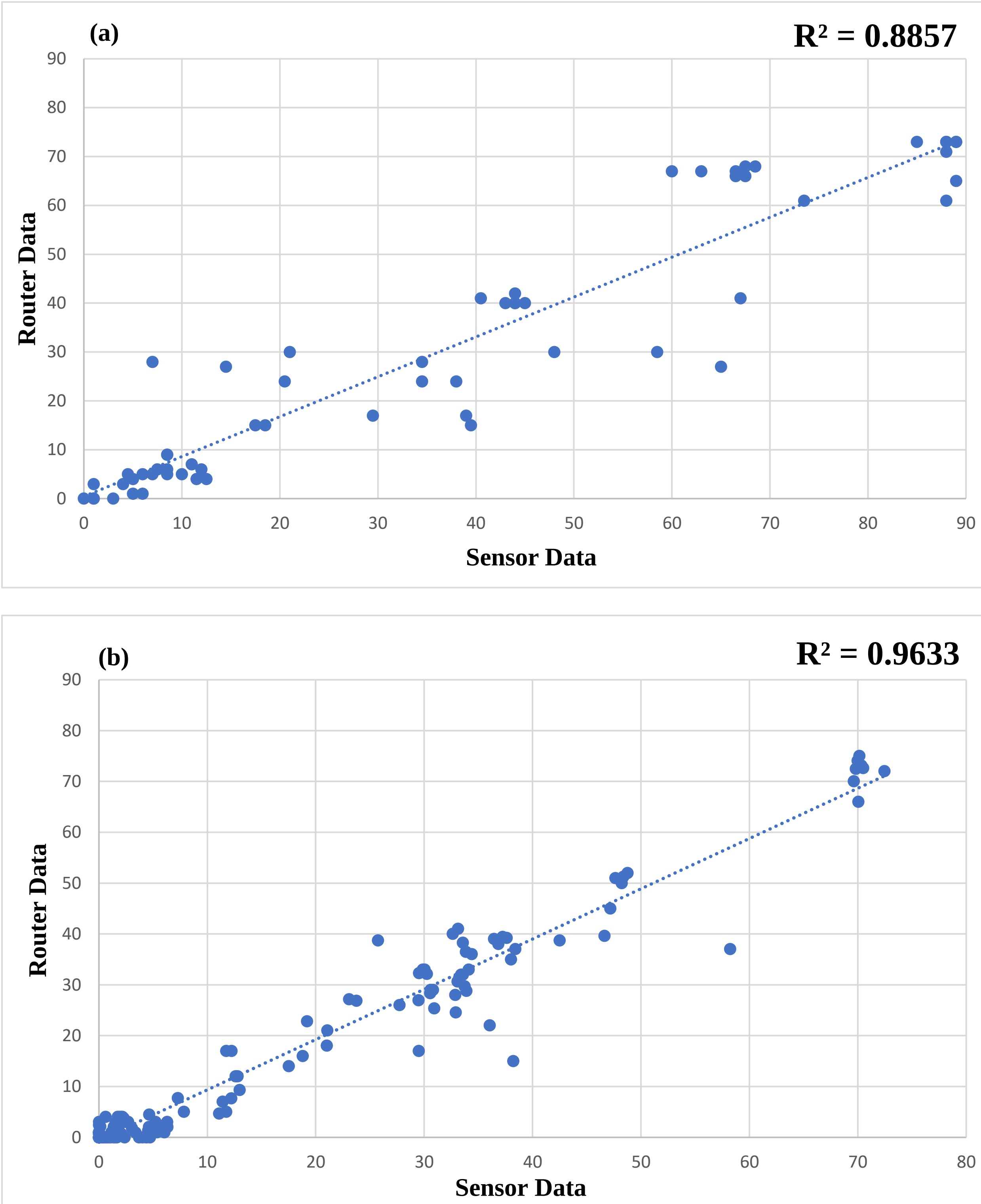
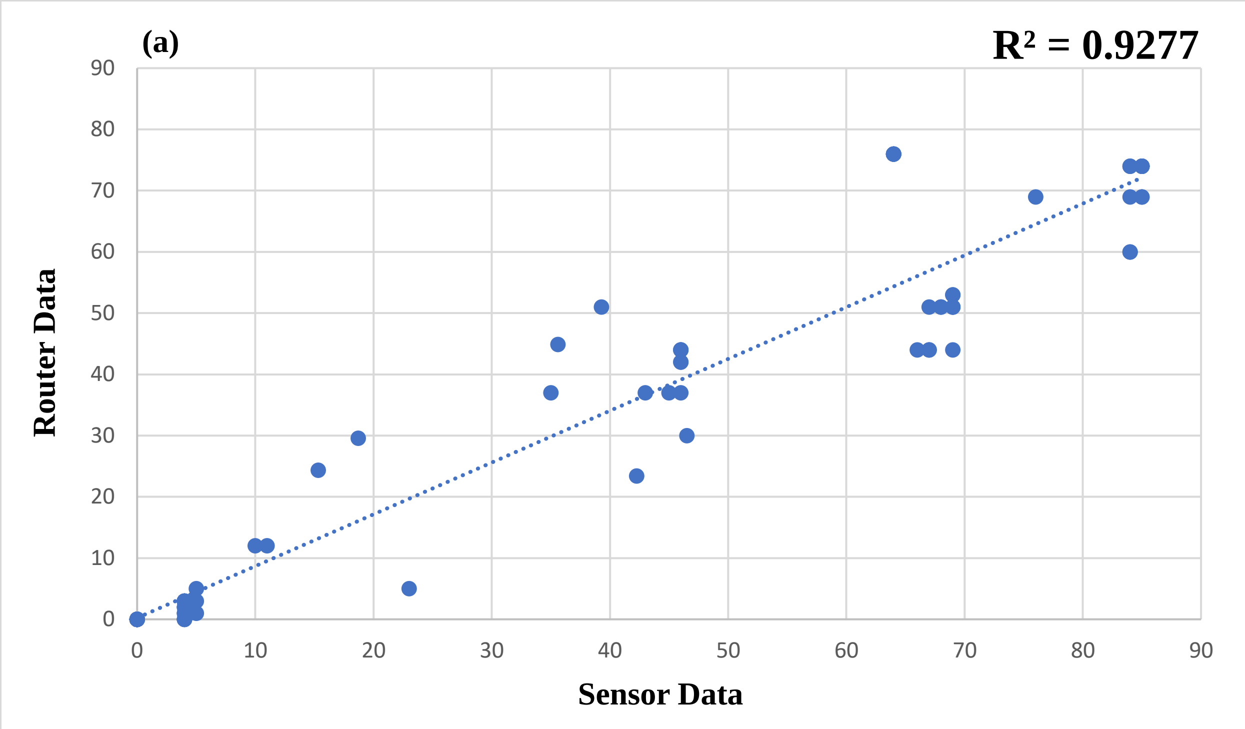


Figure 7: Correlation Plots with Linear Regression Lines. (a) Jan 22, 2019, (b) Jan 23, 2019



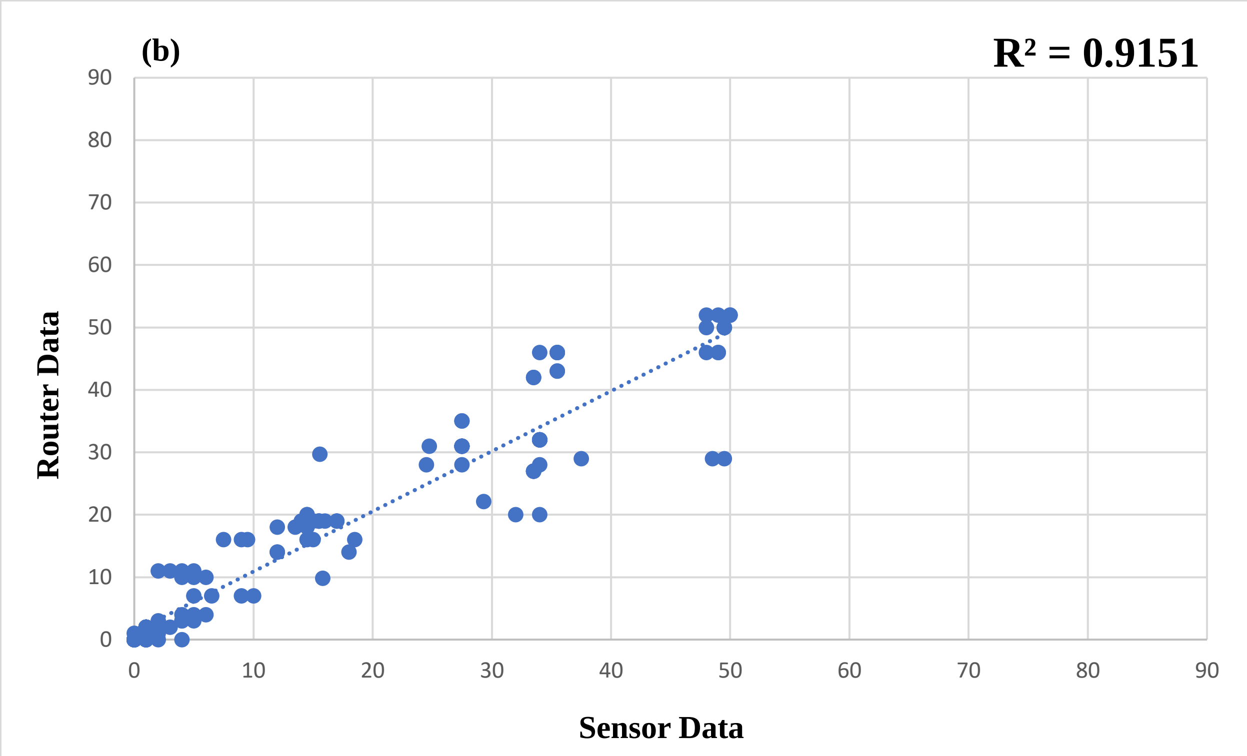


Figure 8: Correlation Plots with Linear Regression Lines: (a) Jan 24, 2019, (b) Jan 25, 2019

Similarly, three weeks (only weekdays) data was analyzed to observe the correlations between the client count from the routers and the occupant count from the C100 sensor. The R^2 values ranged from 0.887 to 0.963 and the intercept of linear regression lines ranged from 0.23 to 1.27. This highlights that the wi-fi router client count agrees with sensor occupancy count. Therefore, it is apparent that meaningful occupancy data can be extracted from the wi-fi routers implying that the routers can serve as an accurate source of occupancy count. Since information that can identify an individual is filtered out by the IT department before the data was pulled out of the servers, the router count has little privacy concerns. Apart from a little cleaning of the raw data, no processing or computation was required to obtain the occupancy count from the routers. This allows the reallocation of computational resources elsewhere while using the router occupancy data for optimizing HVAC's energy usage.

No infrastructure or firmware upgrades were made to the routers to extract the occupancy count from the routers. The student behavior was not altered in anyway during the testing and calibrating of the C100 sensor or during the wi-fi router data acquisition period making this a non-intrusive occupancy estimation technique. The EBTRON C100 sensors costed $450 each and most classrooms in the Mechanical Engineering building have two entrances. Installation of these sensors for every classroom to estimate occupancy is not economically viable. Gathering reliable and accurate occupancy estimates from the w-fi routers can be cost-effective compared to methods that need additional infrastructure, firmware updates, and special operating systems.

4.Conclusion, limitations, and future directions

The R^2 values of 0.887 to 0.963 and linear regression intercept values of 0.23 to 1.27 demonstrate that accurate occupancy counts can be obtained from wi-fi routers with low privacy concerns and minimal computational efforts. As no infrastructure or firmware upgrades were made to the original existing infrastructure, this method has no additional cost impacts. These results address the question raised in the introduction of this paper that wi-fi routers can server as a cost-effective, reliable, and accurate source of occupancy data. However, there are few limitations to this study that need to be addressed in future. The client count from the router may not necessarily represent the total occupants in the lecture hall. There might be instances where students carry more than one wi-fi capable device which may result in over counting of occupants. The wi-fi data needs to be analyzed over many weeks to conclude that routers can provide accurate occupancy counts. These limitations will be addressed in the future steps of this research

References

Azar, Elie, and Carol C. Menassa. 2012. "Agent-Based Modeling of Occupants and Their Impact on Energy Use in Commercial Buildings." *Journal of Computing in Civil Engineering* 26 (4): 506–18. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000158.

———. 2014. "A Comprehensive Framework to Quantify Energy Savings Potential from Improved Operations of Commercial Building Stocks." *Energy Policy* 67: 459–72. https://doi.org/10.1016/j.enpol.2013.12.031.

Chen, Dong, Sean Barker, Adarsh Subbaswamy, David Irwin, and Prashant Shenoy. 2013. "Non-Intrusive Occupancy Monitoring Using Smart Meters." In *Proceedings of the 5th ACM*

Workshop on Embedded Systems For Energy-Efficient Buildings - BuildSys'13. https://doi.org/10.1145/2528282.2528294.

Chen, Zhenghua, Chaoyang Jiang, and Lihua Xie. 2018. "Building Occupancy Estimation and Detection: A Review." *Energy and Buildings* 169: 260–70. https://doi.org/10.1016/j.enbuild.2018.03.084.

Choi, Jeong Woo, Xuanjun Quan, and Sung Ho Cho. 2018. "Bi-Directional Passing People Counting System Based on IR-UWB Radar Sensors." *IEEE Internet of Things Journal* 5 (2): 512–22. https://doi.org/10.1109/JIOT.2017.2714181.

Cisco. 2013. "Location Analytics," 29.

Depatla, S, A Muralidharan, and Y Mostofi. 2015. "Occupancy Estimation Using Only WiFi Power Measurements." *IEEE Journal on Selected Areas in Communications* 33 (7): 1381–93. https://doi.org/10.1109/JSAC.2015.2430272.

Dodier, Robert H., Gregor P. Henze, Dale K. Tiller, and Xin Guo. 2006. "Building Occupancy Detection through Sensor Belief Networks." *Energy and Buildings*. https://doi.org/10.1016/j.enbuild.2005.12.001.

Duarte, Carlos, Kevin Van Den Wymelenberg, and Craig Rieger. 2013. "Revealing Occupancy Patterns in an Office Building through the Use of Occupancy Sensor Data." *Energy and Buildings* 67: 587–95. https://doi.org/10.1016/j.enbuild.2013.08.062.

Hong, Tianzhen, Hongsan Sun, Yixing Chen, Sarah C. Taylor-Lange, and Da Yan. 2016. "An Occupant Behavior Modeling Tool for Co-Simulation." *Energy and Buildings* 117: 272–81. https://doi.org/10.1016/j.enbuild.2015.10.033.

Kleiminger, Wilhelm, Christian Beckel, Thorsten Staake, and Silvia Santini. 2013. "Occupancy Detection from Electricity Consumption Data." In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings - BuildSys'13*. https://doi.org/10.1145/2528282.2528295.

Labeodan, Timilehin, Wim Zeiler, Gert Boxem, and Yang Zhao. 2015. "Occupancy Measurement in Commercial Office Buildings for Demand-Driven Control Applications - A Survey and Detection System Evaluation." *Energy and Buildings* 93: 303–14. https://doi.org/10.1016/j.enbuild.2015.02.028.

Li, Nan, Gulben Calis, and Burcin Becerik-Gerber. 2012. "Measuring and Monitoring Occupancy with an RFID Based System for Demand-Driven HVAC Operations." *Automation in Construction*. https://doi.org/10.1016/j.autcon.2012.02.013.

Liu, Dixin, Xiaohong Guan, Youtian Du, and Qianchuan Zhao. 2013. "Measuring Indoor Occupancy in Intelligent Buildings Using the Fusion of Vision Sensors." *Measurement Science and Technology*. https://doi.org/10.1088/0957-0233/24/7/074023.

Mahdavi, Ardeshir, and Farhang Tahmasebi. 2015. "Predicting People's Presence in Buildings: An Empirically Based Model Performance Analysis." *Energy and Buildings* 86: 349–55. https://doi.org/10.1016/j.enbuild.2014.10.027.

Mantha, Bharadwaj R K, Chen Feng, Carol C Menassa, and Vineet R Kamat. 2015. "Real-Time Building Energy and Comfort Parameter Data Collection Using Mobile Indoor Robots." *International Association for Automation and Robotics in Construction (IAARC)* Proceeding: 1–9. http://www.iaarc.org/publications/fulltext/FFACE-ISARC15-2999958.pdf.

Pan, Shijia, Amelie Bonde, Jie Jing, Lin Zhang, Pei Zhang, and Hae Young Noh. 2014. "BOES: Building Occupancy Estimation System Using Sparse Ambient Vibration Monitoring." *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, 90611O. https://doi.org/10.1117/12.2046510.

SAIC. 2011. "U.S. Smart Grid Case Studies." http://www.eia.gov/analysis/studies/electricity/pdf/sg_case_studies.pdf.

Stoppel, Christopher M., and Fernanda Leite. 2014. "Integrating Probabilistic Methods for Describing Occupant Presence with Building Energy Simulation Models." *Energy and Buildings* 68 (PARTA): 99–107. https://doi.org/10.1016/j.enbuild.2013.08.042.

Vattapparamban, Edwin, Bekir Sait Çiftler, Ismail Güvenç, Kemal Akkaya, and Abdullah Kadri. 2016. "Indoor Occupancy Tracking in Smart Buildings Using Passive Sniffing of Probe Requests." *2016 IEEE International Conference on Communications Workshops, ICC 2016*, 38–44. https://doi.org/10.1109/ICCW.2016.7503761.

Yang, Rui, and Lingfeng Wang. 2013. "Development of Multi-Agent System for Building Energy and Comfort Management Based on Occupant Behaviors." *Energy and Buildings* 56 (January): 1–7. https://doi.org/10.1016/j.enbuild.2012.10.025.

Zou, Han, Hao Jiang, Jianfei Yang, Lihua Xie, and Costas J. Spanos. 2017. "Non-Intrusive Occupancy Sensing in Commercial Buildings." *Energy and Buildings* 154 (September): 633–43. https://doi.org/10.1016/j.enbuild.2017.08.045.

Zou, Han, Yuxun Zhou, Hao Jiang, Szu Cheng Chien, Lihua Xie, and Costas J. Spanos. 2018. "WinLight: A WiFi-Based Occupancy-Driven Lighting Control System for Smart Building." *Energy and Buildings* 158 (November): 924–38.

https://doi.org/10.1016/j.enbuild.2017.09.001.

Zou, Han, Yuxun Zhou, Jianfei Yang, Weixi Gu, Lihua Xie, and Costas Spanos. 2017. “FreeDetector: Device-Free Occupancy Detection with Commodity WiFi.” *2017 IEEE International Conference on Sensing, Communication and Networking, SECON Workshops 2017*. https://doi.org/10.1109/SECONW.2017.8011040.

———. 2018. “FreeCount: Device-Free Crowd Counting with Commodity WiFi.” *2017 IEEE Global Communications Conference, GLOBECOM 2017 - Proceedings* 2018–Janua: 1–6. https://doi.org/10.1109/GLOCOM.2017.8255034.

Zou, Han, Yuxun Zhou, Jianfei Yang, and Costas J. Spanos. 2018. “Device-Free Occupancy Detection and Crowd Counting in Smart Buildings with WiFi-Enabled IoT.” *Energy and Buildings* 174: 309–22. https://doi.org/10.1016/j.enbuild.2018.06.040.

Chapter-03 – Real-Time Occupancy Estimation Using WiFi Network to Optimize HVAC Operation

Krishna Chaitanya Jagadeesh Simma[a], Andrea Mammoli[b]1, Susan M Bogus[c]

[ac]Department of Civil Engineering, University of New Mexico, Albuquerque, 87131, USA

[b]Department of Mechanical Engineering, University of New Mexico, Albuquerque, 87131, USA

ABSTRACT

Commercial and residential buildings consume about 27% of total energy used in the US, out of which nearly half is consumed by commercial building sector, and it expected to grow in the next 30-year period. Literature suggests that occupancy data may improve the energy consumption of the buildings, especially in HVAC operation. In the past few years, studies came up with various frameworks based on existing infrastructure to estimate occupancy, out of which commodity WiFi gained popularity in detecting, estimating, and tracking occupants within buildings. However, there are concerns with those frameworks such as added infrastructure and computational efforts, upgrades to existing infrastructure, and privacy of occupants. This paper presents a simplistic framework based on commodity WiFi to estimate real time occupancy data without any added infrastructure or upgrades, while protecting the occupant privacy and can produce significant energy reduction in HVAC operation. The framework is tested on a large lecture hall in an institutional building that has multiple classes scheduled. The initial tests showed that the WiFi based occupancy had a 0.96 correlation with the established ground truth. Additionally, the WiFi

* Corresponding author. Tel.: +1 505-277-7666.
E-mail address: mammoli@unm.edu

based occupancy schedule resulted in at least 50% savings in HVAC energy consumption over static schedule.

Peer-review under responsibility of the Conference Program Chairs.

Keywords: Occupancy; WiFi; HVAC; Building Management Systems; Energy.

---

1. Introduction

Commercial and residential buildings consume about 27% of total energy used in the US (U.S. Energy Information Institution, 2018a), out of which nearly half is consumed by commercial building sector and it is expected to grow in the next 30-year period. In both commercial and residential sectors, HVAC (Heating, Ventilation and Cooling) holds major share in terms of energy consumed (U.S. Energy Information Institution, 2018b). HVAC is often known to run on poorly controlled building schedules (Carbon Trust, 2007) highlighting the potential for energy savings. Various building energy management systems such as smart BEMS (Building Energy Management Systems) (Rocha, Siddiqui, & Stadler, 2015), MACS (Multi-agent Control Systems) (Dounis & Caraiscos, 2009), and MACES (Multi-agent comfort and energy system) (Klein et al., 2012) among others were proposed in literature aimed at reducing building energy consumption and improving occupant comfort. Although these frameworks reported potential reductions in building energy consumption, they do not incorporate real-time occupancy information into their systems. The frameworks assume peak-hour, off-peak hour occupant loads, and occupant preferences to program thermostat setpoints. This allows ample room for additional building energy savings that can be achieved when occupant count is used to manage building energy consumption while maintaining occupant comfort.

The occupant impact in reducing building energy consumption is highlighted by Azar and Menassa (Azar & Menassa, 2012c) which signifies the importance of collecting occupant information. The findings of Azar and Menassa (Azar & Menassa, 2012c) led the following studies to develop various techniques to obtain occupancy information that can be used to improve building energy consumption. Infrastructure networks such as cameras, RFID sensors, PIR sensors, and motion sensors among other to detect and estimate occupancy were proposed across studies (Z. Chen et al., 2018). However, costs associated in establishing dedicated networks led researchers to look for cost effective solutions to estimate occupancy in commercial buildings. Most modern buildings are equipped with infrastructure such as CO_2 sensors, temperature sensors, humidity sensors, smart meters, and WiFi networks, among others. These infrastructure networks exhibit variations in their readings in occupant presence allowing for occupant detection, estimation, and tracking contingent on type of infrastructure. Recent studies started evaluating the level of detection and estimation accuracy of these infrastructure networks (Z. Chen et al., 2018). Out of the listed existing infrastructure, occupancy based on WiFi networks gained popularity in the past three years owing to its availability and wide coverage inside the many commercial buildings [4 –11].

The WiFi network facilitates detection, estimation and tracking of users connected to the network within the building through the MAC (Media Access Control) addresses (Vattapparamban et al., 2016). This idea has been implemented in various occupancy estimating frameworks such as WiFree (Zou, Zhou, Yang, & Spanos, 2018), WinOSS (Zou, Jiang, et al., 2017), Wi-Fi Pineapple

(Vattapparamban et al., 2016), and Meraki (Cisco, 2013). These methods capture the received signal strength (RSS) of probe requests using wireless sniffers to infer the occupant count in the test zone. Using the RSS facilitates tracking of a specific user within the range of the WiFi network for added accuracy. Recently introduced MAC address randomization may increase the inaccuracy of the occupancy obtained by scanning probe request (Vattapparamban et al., 2016). Other methods use channel state information (CSI) to estimate the number of occupants by forming transmitter and receiver pairs (TX - RX) [5–7]. For this system to work, set of two routers are needed for the test space where one router acts as transmitter (TX) and the other acts as receiver (RX) (Zou, Zhou, Yang, Gu, et al., 2018). However, these methods require upgrades to the existing infrastructure, additional infrastructure (a second pair of routers), and heavy computational requirements to capture CSI data to infer the occupancy data. Additionally, MAC addresses can be used to identify an individual and track their location and activity over the network which raises the privacy concerns of the occupants. Table 2 summarizes the existing frameworks to detect occupancy using WiFi networks.

Table 2: Existing WiFi based Occupancy Detecting Frameworks

Name	Processing Method	Additional Resources	Type of Test Zone	Concerns	Source
WinOSS	RSS	Firmware upgrades	Controlled space	Occupant identification	(Zou, Jiang, et al., 2017)
WiFree	Channel State Information (CSI)	Second Wi-Fi Router, upgrade, and new firmware	Controlled small space	Computational requirements	(Zou, Zhou, Yang, & Spanos, 2018)
Meraki	MAC address	Meraki wireless APs	Large buildings	Occupant identification	(Cisco, 2013)

Wi-Fi Pineapple	RSS	Wi-Fi Sniffers	Large institutional building	Occupant identification	(Çiftler et al., 2018)(Vattapparamban et al., 2016)
FreeDetector	Channel State Information (CSI)	Firmware upgrades	Controlled space	Computational requirements	(Zou, Zhou, et al., 2017)
FreeCount	Channel State Information (CSI)	Firmware upgrades	Controlled small space	Computational requirements	(Zou, Zhou, Yang, Gu, et al., 2018)

From the table it is evident that existing methods either requires complex information processing or a personal identifier (i.e., MAC address) to obtain occupancy data. Added infrastructure such as secondary router or probe sniffers or firmware upgrades are needed to execute the exiting frameworks which may also hinder large scale deployment in commercial buildings. Most importantly, majority of these frameworks were tested over small controlled spaces with limited occupants where the occupant dynamics does not represent the complexity of a real-time occupant movement in commercial buildings (Zou, Jiang, et al., 2017; Zou, Zhou, Jiang, et al., 2018; Zou, Zhou, Yang, & Spanos, 2018). As high discrepancies were noted between predicted energy consumption and actual consumption (Azar & Menassa, 2012c) due to occupant impact, it is important to acquire real-time occupant data especially in institutional buildings. Institutional buildings such as university campus buildings have large number of occupants entering and exiting throughout the day. For examples, students may use the building for only during their scheduled classes which are an hour to two hour long and leave the building when done. In such scenarios the building can be occupied by many unique occupants within a day which raises the randomness in occupants and occupant behavior. With the increased randomness in occupants and occupant behavior, it is necessary to understand the HVAC energy consumption while improving occupant comfort levels and saving energy.

Using occupancy based operating schedules as opposed to fixed schedules reportedly increase the HVAC energy savings (J. Yang, Santamouris, & Lee, 2016). However, it is vital that the occupancy schedules used to estimate the reported HVAC savings need to represent the dynamics of a real-time occupant movement within the buildings (Kwok, 2011). With increasing number of smart buildings that are equipped with WiFi networks covering the entire building, it is possible to obtain such real-time occupant data that can be used to optimize HVAC operations. To allow building owners to maximize their energy saving with minimal efforts, a framework to estimate accurate real-time occupancy with nominal computational efforts and no additional infrastructure requirements that also addresses the privacy concerns is necessary. As a part of the framework development the initial findings on WiFi occupancy data are presented in (Krishna Chaitanya J Simma, Bogus, & Mammoli, 2019). The current paper summarizes the findings of the initial work and continues developing the framework. In this context, this paper asks a question: *What amount of HVAC energy savings can be achieved using real-time occupancy data from WiFi network in an institutional building?*

2. Methodology

To address the issues highlighted in the introduction and the question, a methodology is proposed as shown in **Figure 9**. The proposed methodology is divided into two phases, Phase-I and Phase-II. In phase-I, the ground truth is established using people counting sensor (EBTRON, C100) and occupancy data are extracted from the WiFi network with the assistance of the university's Information Technology department.

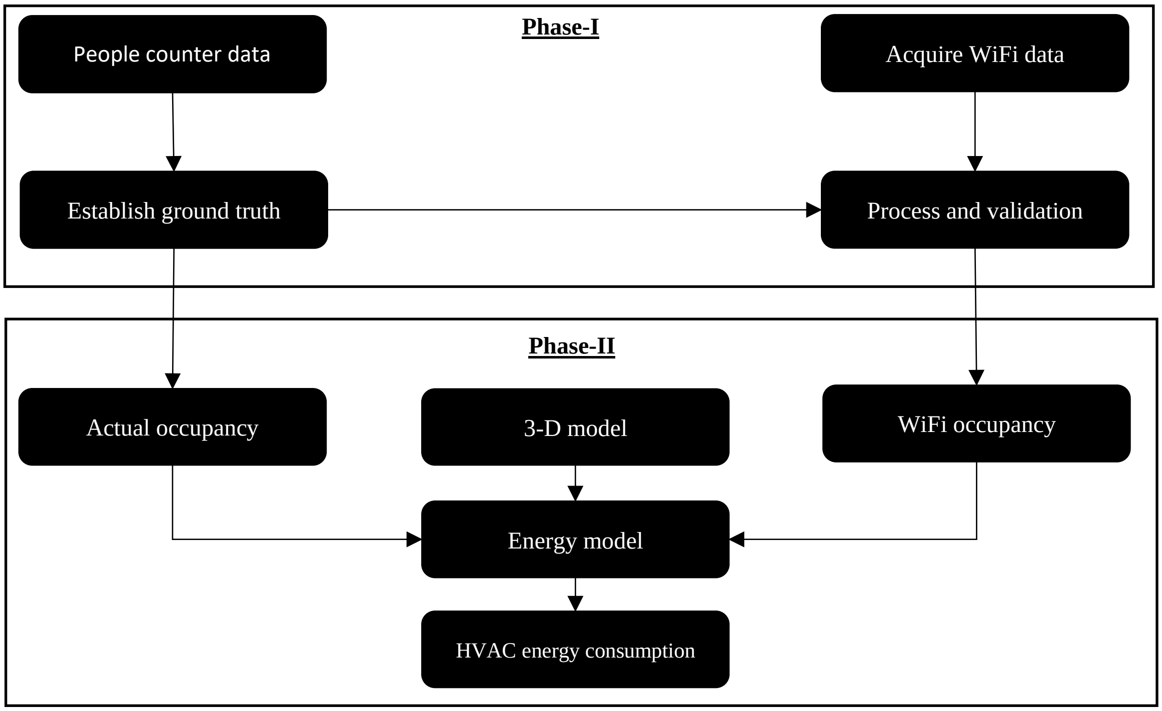


Figure 9: Methodology (Objective-II)

The occupancy data from WiFi network is validated against the ground truth to estimate accuracy. The results of Phase-1 are summarized in this paper but for detailed steps and results on Phase-I, readers are encouraged to refer to (Krishna Chaitanya J Simma et al., 2019). Phase-II consists of following steps: 1) build a 3D model for the area of study, 2) Build energy models and HVAC operating schedules using occupancy data from people counting sensors, WiFi network, and fixed schedule, 3) Perform energy analysis to estimate potential HVAC energy savings using fixed schedules, ground truth and WiFi based occupancy schedules.

A large lecture hall (Rm-218: 26’x40’) inside the thirty-year old Mechanical Engineering at University of New Mexico (UNM) was used for this study. The ME building is equipped with campus WiFi network with routers throughout the building and Rm-218 was preinstalled with three WiFi routers. The underlying assumption in estimating occupancy using WiFi network is that students who enter the building carry at least one WiFi capable device and connects to the campus network during their stay. UNM's Information Technology department provided the data of number of users connected to the network inside the lecture hall.

2.1. Methodology overview

The methodology is aimed to establish a framework that allows large scale implementation. To achieve the set goal, the following criteria were established: 1) *Minimal infrastructure and computational effort*: estimating occupancy is the primary input to optimize the HVAC operation and occupant comfort. It is essential that the occupancy data can be obtained with minimal effort that can be utilized in occupancy predictions via machine learning. 2) *Privacy*: Commercial buildings include office buildings, public libraries, governmental offices, shopping malls, and institutional buildings. These are often occupied by public, and it is of paramount importance that the privacy of the public is protected. 3) *Real-time non-intrusive occupant dynamics:* As institutional buildings tend to have different students occupying a space throughout the data, it is essential to capture the true dynamics of occupant movements throughout the day.

3. Results and Discussion

3.1. Phase-I

Phase-I is aimed at establishing the ground truth and estimating the accuracy of the WiFi based occupancy data. **Figure 9** shows the comparison between the established ground truth (from C100 sensor) and router data obtained from the IT department. The router data comprises of number of clients connected to the campus network inside the Rm-218 with a timestamp. The sensor data is the occupant count obtained from the people counting sensors installed on the door frames which also serves as ground truth.

A single day from each of the first six weeks of data collected is shown in **Figure 10**. The ground truth established by the people counting sensor on an average estimated 97.7% of the occupancy of the room obtained by manual count. The data presented in **Figure 10** shows students entering and leaving the room under study for their scheduled classes throughout the day. The occupant's behavior was not modified in anyway during the data collection duration. Therefore, the plots in **Figure 10** show the natural movement of occupants in and out of the room reflecting a realistic occupant information. The maximum number of users connected to the WiFi routers in the same room seem to follow the same trend as the occupant count. The WiFi router data when compared to the ground truth achieved 96% accuracy. The correlation plot for occupant count and maximum number of clients connected and the R^2 values for the days shown in **Figure 10** were presented in **Figure 11**. The R^2 values ranged between 0.86 to 0.96 for the first seven weeks of the data collected indicating strong correlation between ground truth and the WiFi occupant count.

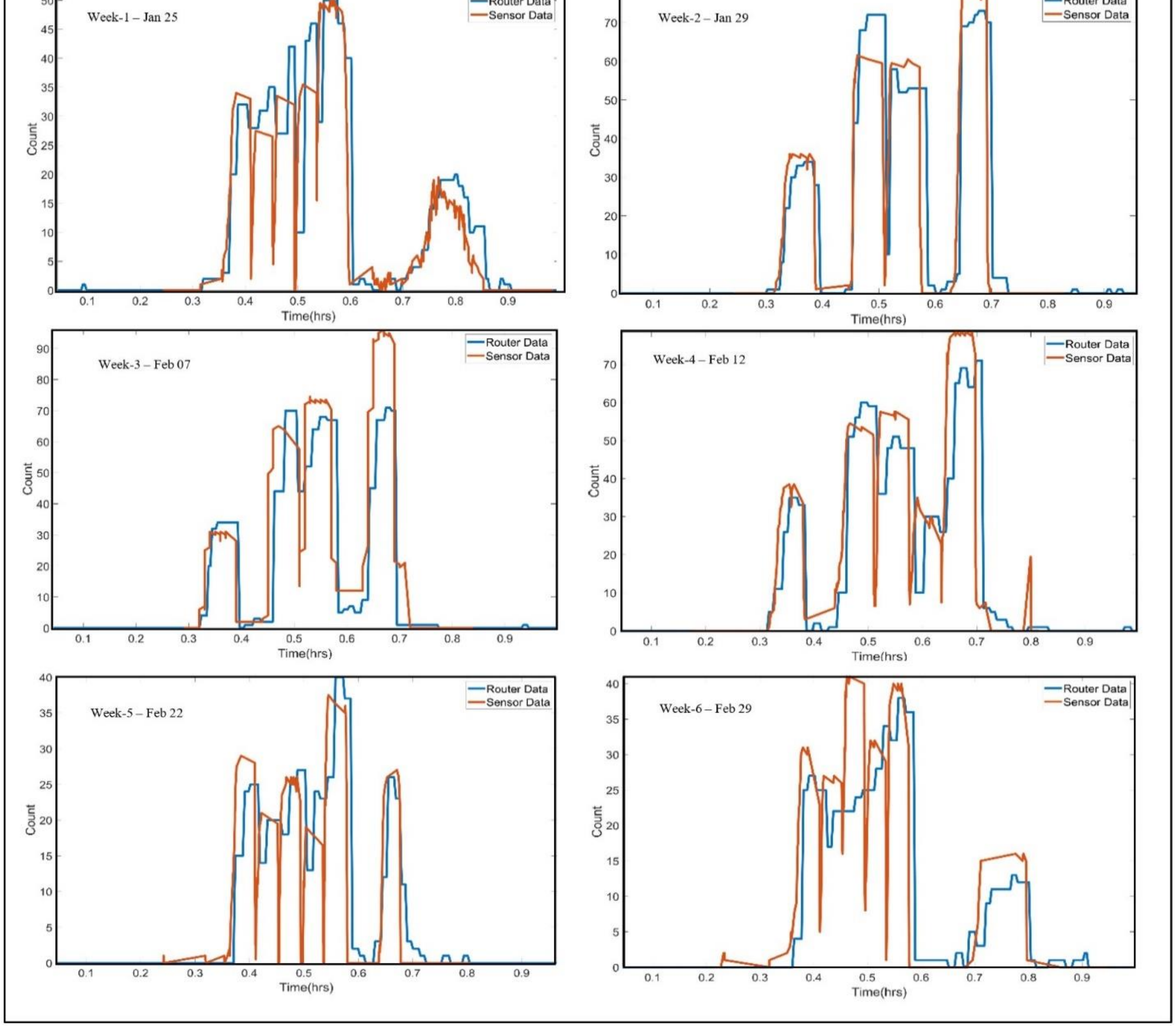


Figure 10: Occupancy data comparison for a single day from each of the first six weeks of the study period

3.2. Phase-II

To validate the utility of WiFi occupancy data further, an energy analysis needs to be performed on HVAC operation to ensure that the WiFi data results in significant energy savings. To perform building energy analysis with occupancy data, a set of software tools are required to interoperate in collaboration. For this study, the software used to build the energy models are SketchUp - a 3D modeling tool developed by Trimble, OpenStudio - a cross-platform tool set capable of running whole building energy models, and EnergyPlus - a DOE (Department of Energy) developed building energy simulation program.

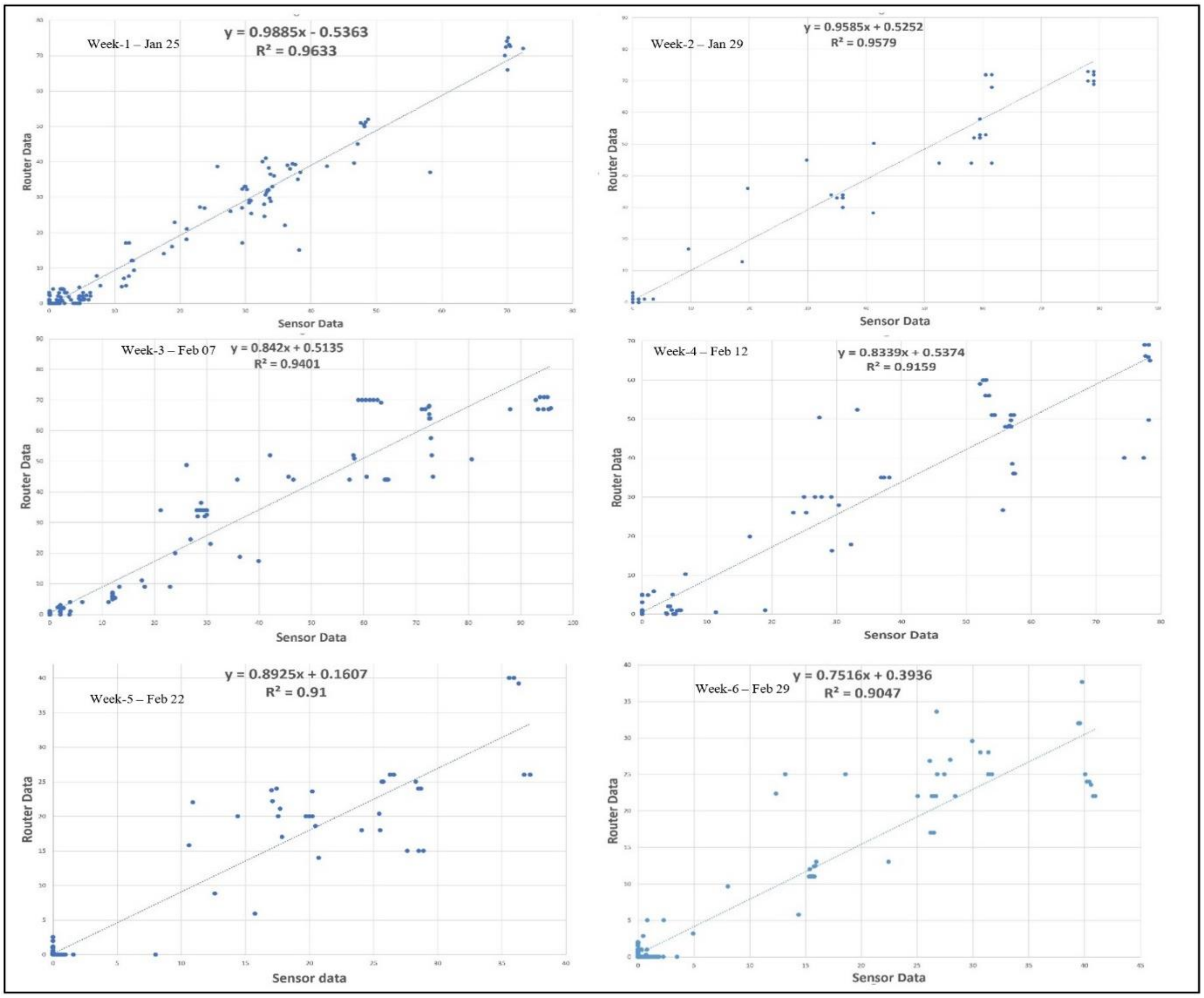


Figure 11: Correlation plots for router data versus people counter data

For a comparative analysis, occupancy schedules were built using ground truth (C100 sensor data), WiFi data, and fixed schedule. The energy consumption associated with ground truth and WiFi occupant schedules were compared against the fixed schedule. Most of the buildings operate on a fixed schedule (specific time period in a day) assuming maximum occupancy (Z. Yang & Becerik-Gerber, 2014). Using the fixed schedule as a baseline can help understand the maximum amount of savings that can be achieved using occupant data. Additionally, by comparing the occupant schedule derived by WiFi data to the baseline and ground truth can validate the accuracy of the WiFi data.

3.1.1. Building modelling

The lecture hall used in this study is shown in **Figure 12** (a) and (b). **Figure 12** (c) shows the lecture hall layout. A 3D model is built using SketchUp Pro-09 which creates wall surfaces with thermal properties, location details, and shading areas. A legacy OpenStudio plugin version 0.9.3 is needed for SketchUp to convert the building model to energy model shown in **Figure 12** (d).

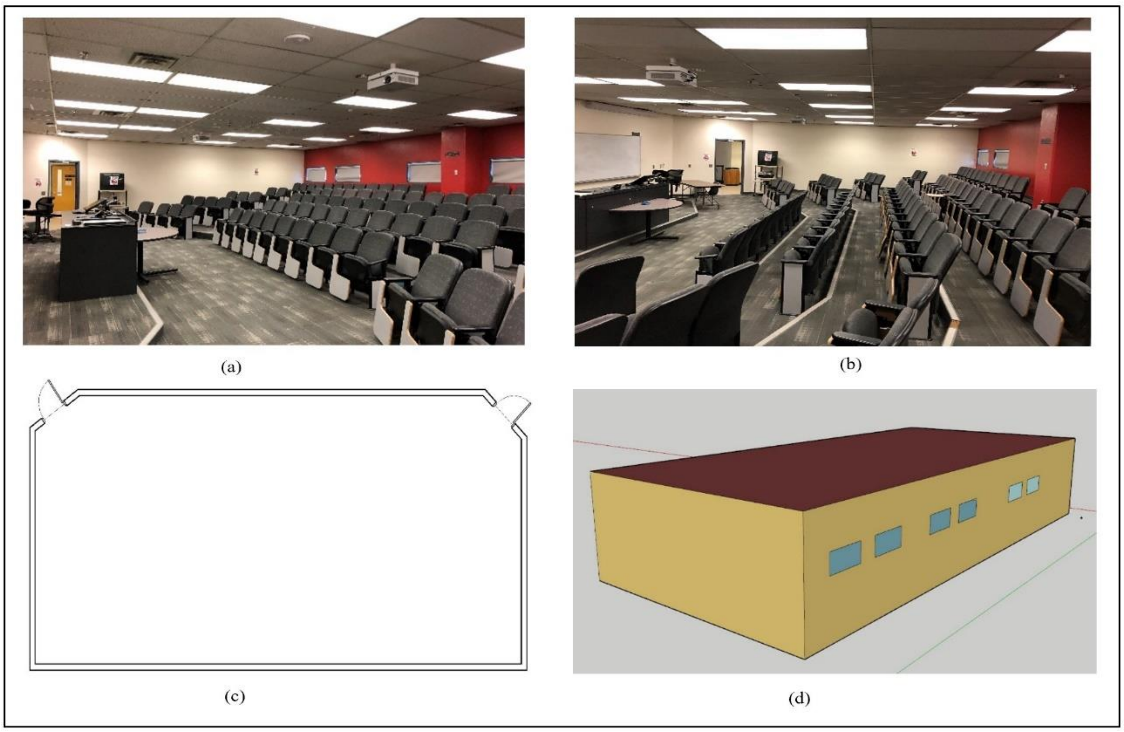


Figure 12: (a) & (b) Lecture Hall, (c) Lecture Hall layout, (d) Energy Model for the lecture hall

3.1.2. Energy modelling and occupancy schedules

The energy model file created in SketchUp is exported as an EnergyPlus input data file (IDF) that can be opened by EP-Launch (EnergyPlus v9.1.0). A pre-designed weather file for the location is added to run the energy model. The programs allow for modeling different schedules for heating, cooling, office equipment, lights, and occupancy. Using the ground truth and WiFi data occupancy schedules and heating schedules were created for energy models. Since the testing period is in fourth week of January, only heating schedules were created. The fixed schedule is designed to operate HVAC between from 7am to 7pm at maximum occupant capacity. The simulation is run only for the first week of the data collected. The models were run for similar set of parameters while the occupancy schedules were altered in each run. This allows to capture the energy consumption changes due to occupant data variations.

3.1.3. HVAC Energy analysis

As explained in section 0 three occupancy schedules were created for 1) fixed schedule, 2) ground truth schedule, and 3) WiFi data schedule and the fixed schedule is used as a baseline to compare the rest of the schedules. The HVAC energy consumption was estimated for each schedule from the model are presented in Table 3. As expected, the fixed schedule energy consumption for HVAC is the highest of the three schedules modelled. The lowest energy consumption was noted by the ground truth schedule and the WiFi schedule was also closer to the ground truth schedule.

The ground-truth schedule has only 3% lower energy consumption than WiFi schedule indicating slight difference and 57% lower than fixed schedule. The WiFi schedule’s energy consumption was 56% lower than the energy consumption of the fixed schedule. However, setting up people counting sensors for each classroom is expensive. In this case, WiFi routers function as people counting sensors which does not affect HVAC stability as WiFi data is used for occupancy forecast only while regular temperature sensors maintain the temperature of the zone. Additionally, WiFi data provided with near identical results to the ground truth data highlighting that a simple client count data from the routers can result in significant HVAC energy savings.

Table 3: Energy consumption data from energy models

Schedule Type	HVAC Energy Consumption [MJ/m^2]	Energy difference [MJ/m^2]	Percentage savings
Fixed Schedule	591.70	-	-
Ground truth occupancy	253.70	338.53	57%
WiFi occupancy	260.93	330.77	56%

4. Conclusion

It is well established that demand control HVAC operation can result in significant energy savings. Obtaining occupancy data has become a major concern for commercial buildings to operate HVAC. The expense of establishing dedicated infrastructure to obtain occupancy data may lead to building owners shying away from taking steps towards energy conservation. Therefore, this study proposed a WiFi based occupancy data collection with minimal computational efforts and limited privacy concerns. The WiFi data obtained from UNM's IT department consists of client count with a timestamp for each day. This data consists of no identifiers such as MAC addresses of the users connected to the network and thus it minimizes the privacy concerns of the occupants. The data is received as a CSV (comma separated value) file from the IT department which requires minimal denoising and thus reducing the computational requirements. The WiFi data achieved highest correlation of .964 and R2 values for the next six weeks of data collected stayed between 0.86 to 0.96.

The energy simulations using occupancy data from people counting sensor and WiFi routers resulted in over 50% lower HVAC energy consumptions compared to the fixed schedule. The sensor schedule resulted in energy consumption similar to that of the WiFi schedule with only 3% variation. However, the costs associated in establishing the ground truth hinders its wide range application. This highlights that WiFi data can achieve HVAC energy savings that are similar to the savings achieved by actual occupancy data without any added infrastructure or costs. Additionally, the occupant behavior was not altered in anyway during the data collection for both sensor data and WiFi data thus capturing the natural movement of the occupants in Rm-218. Therefore, the proposed methodology successfully captured the natural dynamics of the occupant movements with minimal expenses and computation. Furthermore, the methodology avoids the collection of any personal identifiers thereby reducing the privacy concerns of the occupants.

Hence reliable and accurate occupant data can be obtained using simple client count which can also result in significant energy savings in HVAC operation. The 56% lower HVAC energy consumption using WiFi schedule answers the question asked in section 1. Saving more than half the energy consumed by a fixed schedule provides confidence to move the current study forward to the next phases of this study. However, these savings can only be observed in buildings which are occupied only during specific hours of the day such as schools, university buildings, public libraries, and office spaces with specific working hours. Buildings that are occupied throughout the day may not result in reported savings. With preliminary results leading to solid energy savings, more data need to be analyzed to ensure the patterns follow over longer periods.

References

[1] U.S. Energy Information Institution, “Annual Energy Outlook 2018 with projections to 2050,” *Annu. Energy Outlook 2018 with Proj. to 2050,* vol. 44, no. 8, pp. 1–64, 2018.

[2] Carbon Trust, “CTV032: Technology Overview: Building controls,” 2007.

[3] P. Rocha, A. Siddiqui, and M. Stadler, “Improving energy efficiency via smart building energy management systems: A comparison with policy measures,” *Energy Build.*, vol. 88, pp. 203–213, 2015.

[4] A. I. Dounis and C. Caraiscos, “Advanced control systems engineering for energy and comfort management in a building environment-A review,” *Renew. Sustain. Energy Rev.*, vol. 13, no. 6–7, pp. 1246–1261, 2009.

[5] L. Klein *et al.*, “Coordinating occupant behavior for building energy and comfort management using multi-agent systems,” *Autom. Constr.*, vol. 22, pp. 525–536, 2012.

[6] E. Azar and C. C. Menassa, “Agent-Based Modeling of Occupants and Their Impact on Energy Use in Commercial Buildings,” *J. Comput. Civ. Eng.*, vol. 26, no. 4, pp. 506–518, Jul. 2012.

[7] Z. Chen, C. Jiang, and L. Xie, “Building occupancy estimation and detection: A review,” *Energy Build.*, vol. 169, pp. 260–270, 2018.

[8] H. Zou, Y. Zhou, J. Yang, W. Gu, L. Xie, and C. Spanos, “FreeCount: Device-free crowd counting with commodity WiFi,” *2017 IEEE Glob.*

Commun. Conf. GLOBECOM 2017 - Proc., vol. 2018–Janua, pp. 1–6, 2018.

[9] S. Depatla, A. Muralidharan, and Y. Mostofi, “Occupancy Estimation Using Only WiFi Power Measurements,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1381–1393, 2015.

[10] H. Zou, Y. Zhou, J. Yang, W. Gu, L. Xie, and C. Spanos, “FreeDetector: Device-Free Occupancy Detection with Commodity WiFi,” *2017 IEEE Int. Conf. Sensing, Commun. Networking, SECON Work. 2017*, 2017.

[11] H. Zou, Y. Zhou, J. Yang, and C. J. Spanos, “Device-free occupancy detection and crowd counting in smart buildings with WiFi-enabled IoT,” *Energy Build.*, vol. 174, pp. 309–322, 2018.

[12] B. S. Çiftler, S. Dikmese, I. Güvenç, K. Akkaya, and A. Kadri, “Occupancy Counting with Burst and Intermittent Signals in Smart Buildings,” *IEEE Internet Things J.*, vol. 5, no. 2, pp. 724–735, 2018.

[13] H. Zou, H. Jiang, J. Yang, L. Xie, and C. J. Spanos, “Non-intrusive occupancy sensing in commercial buildings,” *Energy Build.*, vol. 154, no. September, pp. 633–643, 2017.

[14] E. Vattapparamban, B. S. Çiftler, I. Güvenç, K. Akkaya, and A. Kadri, “Indoor occupancy tracking in smart buildings using passive sniffing of probe requests,” *2016 IEEE Int. Conf. Commun. Work. ICC 2016*, pp. 38–44, 2016.

[15] Cisco, “Location Analytics,” p. 29, 2013.

[16] H. Zou, Y. Zhou, H. Jiang, S. C. Chien, L. Xie, and C. J. Spanos, “WinLight: A WiFi-based occupancy-driven lighting control system for smart building,” *Energy Build.*, vol. 158, no. November, pp. 924–938, 2018.

[17] J. Yang, M. Santamouris, and S. E. Lee, “Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings,” *Energy Build.*, vol. 121, pp. 344–349, 2016.

[18] S. S. K. Kwok, “A study of the importance of occupancy to building cooling load in prediction by intelligent approach,” *Energy Convers. Manag.*, vol. 52, no. 7, pp. 2555–2564, 2011.

[19] K. C. J. Simma, S. M. Bogus, and A. Mammoli, “WI-FI ROUTER NETWORK-BASED OCCUPANCY ESTIMATION TO OPTIMIZE HVAC ENERGY CONSUMPTION,” vol. 2015, no. Saic 2011, pp. 1–10, 2019.

[20] Z. Yang and B. Becerik-Gerber, “The coupled effects of personalized occupancy profile-based HVAC schedules and room reassignment on building energy use,” *Energy Build.*, vol. 78, pp. 113–122, 2014.

Chapter-04 - Fuzzy ART: Pattern Recognition of WiFi Detected Occupancy in Commercial Buildings

Krishna Chaitanya Jagadeesh Simma[1], Thomas P. Caudell[2], Susan M Bogus[3], Andrea Mammoli[4.]

[1]PhD candidate, Department of Civil, Construction and Environmental Engineering, University of New Mexico, MSC01 1070, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: jagadeesh145@unm.edu

[2]Emeritus Professor, Department of Electrical and Computer Engineering, University of New Mexico, MSC01 1070, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: tcaduell@unm.edu

[3]Professor, Department of Civil, Construction and Environmental Engineering, University of New Mexico, MSC01 1070, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: sbogus@unm.edu

[4]Customer Technologies Program, Electric Power Research Institute; 3420 Hillview Ave, Palo Alto, CA 94304; e-mail: amammoli@epri.com

Abstract

Research that predicts occupancy patterns in commercial buildings has gained in significance ever since the influence of occupants on building energy consumption became evident. Studies have employed a variety of sensory systems to collect the occupancy data and understand human behaviors throughout buildings. However, establishing a dedicated sensor network to collect occupancy data can become expensive for building owners. In this context, obtaining occupancy data from an existing WiFi network could eliminate the cost concerns. Data within the WiFi routers

provide sufficient information for accurate estimates of occupancy. To estimate occupancy levels, this work proposes to learn and recognize WiFi connection using an Adaptive Resonance Theory (ART) artificial neural network. A detailed understanding of occupancy patterns using the WiFi data is helpful for developing heating and cooling schedules that optimize HVAC energy consumption. For this study, occupancy data was collected over a 17-week semester at the University of New Mexico using existing WiFi routers located in a large lecture hall used by multiple classes. This data was used to learn patterns of repetition using the neural network. The results show that if the 24-hour occupancy profiles can be subdivided into smaller time segments defined by external schedules such as lecture start and end times or other constraints, significant patterns can be detected. A detailed understanding of these patterns can greatly facilitate occupancy load forecasting for effective building management (e.g., HVAC operation).

1. Introduction

The U.S. Energy Information Administration (EIA) projects global building energy consumption to increase by 1.3% per year for the next three decades (EIA & Hojjati, 2019). In the United States, commercial building energy consumption has increased over the past four decades and is projected to increase by 0.5% per year for the next three decades while residential consumption is projected to decrease by 0.1% per year during the similar time period (U.S. Energy Information Institution, 2018a). Similarly, emerging economies like India, China, and other Southeast Asian and African countries are expected to increase their building energy consumption by 2% annually for the next three decades (EIA & Hojjati, 2019). China’s building energy consumption increased 1.7 times from 2001 to 2014 and it is over 20% of its current energy consumption (Huo et al., 2020). Similarly, India has doubled its building energy consumption since 2000 and it is over 33% of its total energy consumption (Basu et al. 2014). India’s energy consumption is expected to grow at a rate of 2.7% until 2040, higher than any other region (Outlook, 2017). Additionally, India is expected to increase its commercial floor space by 13,000 million square feet and the US is expected to increase almost twice that amount by 2030 (Basu et al. 2014). Despite the slowdown since 2015, China is also expected to increase its commercial floor space in the next two decades (Jiang, Yan, Guo, & Hu, 2019).

From these projections, it is evident that global commercial building energy demand will increase for the next three decades. Over the past decade, studies have established the critical role of occupancy in commercial building energy consumption (Chen et al. 2018). In this context, numerous frameworks were proposed to detect, estimate, and track occupants within buildings (Chen et al. 2018). To this extent, a basic framework to estimate occupancy from WiFi routers was proposed in (Simma et al. 2019). Preliminary energy simulations showed that significant HVAC energy savings can be achieved by using occupancy detected by the proposed framework (Simma et al. 2019). However, occupancy in commercial buildings such as airports, libraries, gymnasiums, schools, colleges, and universities tend to change over time. To establish the consistency of potential energy savings reported in (Simma et al. 2019), it is imperative to understand the variety of occupancy patterns and its impact on building energy consumption.

In the past decade, studies proposed frameworks that utilize the data collected from various sensory networks within the buildings to measure the occupant patterns for building energy analysis. The occupant behavior models can be generalized into deterministic models that use diversity in occupant behavior (Ding, Chen, Wei, & Yang, 2021), stochastic models such as Agent-based models and Markov chains (Gaetani et al. 2016) and machine learning models such as Support vector machine (SVM), K-nearest neighbors (kNN), and Artificial Neural Networks (ANN) (Wang et al. 2018). In general, the deterministic and stochastic models (e.g., Monte Carlo) focused on predicting specific individual behaviors such as opening of windows, user preferences (e.g., lights, blinds, heaters, and fans), activities and clothing. While the machine learning models (e.g., ANN, kNN, SVM) focused on predicting occupancy profiles for single/multi office spaces with small occupancy volumes.

In W. Wang et al. (2018) experiments of occupancy prediction, the ANN model reportedly had higher accuracy over the kNN and SVM models. Additionally, the accuracy and reliability of these models can be increased by using fused data from multiple sensor networks (e.g., WiFi probe data (MAC/IP address) fused with environmental sensors). On the other hand, the Hidden Markov model (HMM) is also widely used in occupancy predictions (Dong et al. 2010, Ding et al. 2021). Studies that implemented various forms of HMM reported occupancy prediction accuracy ranging from 73% to 98.4% (Ding et al., 2021). A summary of some of the occupant prediction studies along with the description of test space used and occupant volume is presented in **Table 4**.

Table 4: Summary of Occupant Prediction Studies

(Ding et al., 2021)	K-means	Occupancy state (in or out)	Multiple university buildings	Variable
(Wang, et al. 2018)	ANN, SVM, kNN	Occupancy number	Office	18
(B. Dong et al., 2010)	HMM	Occupancy number	Open-plan Office	4
(Liao et al.2012)	Agent-based	Occupancy	Office	Single and multi-occupant scenarios
(Parys, et al.2011)	Markov-Chain	User Behavior	Office	-

From **Table 4** it is evident that these studies were conducted in small spaces (e.g., single/multiple office) with limited occupancy. These spaces do not represent the dynamic time variant occupancy patterns of large commercial spaces such as airports, public libraries, shopping centers, university buildings and gymnasiums, among others. Large commercial spaces have significantly high variability in occupancy compared to single/multi office spaces. Additionally, these spaces might not be occupied by the same set of occupants (e.g., shoppers at shopping malls, and travelers at airports). Current studies focus on understanding occupant behavioral aspects (e.g., opening of windows, lighting, and activities) to reduce energy consumption. The strategies to reduce energy consumption in these studies are purely based on occupant behavior modification. However, in large commercial spaces with dynamic occupancy, individual behaviors patterns such as temperature preferences, opening windows, and lighting preferences among others might have little impact on the overall building energy consumption since individuals do not have the ability to modify settings for personal preference. A methodology to detect occupancy patterns in complex occupancy environments such as airports, libraries, universities, and public buildings is needed.

From the occupancy data collected in previous studies (Simma et al. 2019, Simma et al. 2019), it was evident that university lecture halls follow specific occupancy schedules for all weekdays. This implies that days with similar scheduled occupancy follow a specific pattern. Detecting these patterns can aid in occupancy load forecast studies and occupant comfort studies. To this extent, Fuzzy Adaptive Resonance Theory (ART) (Carpenter et al. 1991) was used to analyze occupancy data collected from WiFi routers for a large lecture hall at University of New Mexico. Fuzzy ART is a neural network approach that is capable of rapid learning and representing categories of patterns in data. Additionally, Fuzzy ART's encoding of its inputs guarantees single pass convergence during the learning process. The lecture hall in this study has a capacity of 100 occupants with occupancy varying between 30% to 90% of total capacity depending on the lecture scheduled, thus, providing a complex test environment where occupancy has regularity over daily and weekly time scales. Therefore, it is hypothesized that Fuzzy ART detects pattern in the occupancy profiles from a complex environment and clusters them in categories.

2. Methodology

The methodology shown in Figure 13 is used to identify occupancy patterns in a large lecture hall using WiFi routers: 1) Data acquisition, 2) Data normalization, 3) Fuzzy ART clusters, 4) Output processing and analysis.

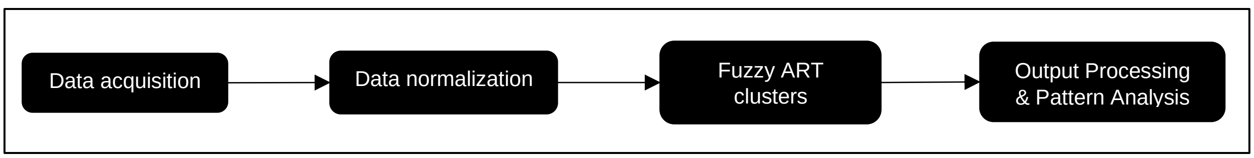


Figure 13: Methodology (Objective-III)

2.1.Data Acquisition

The occupancy data for the lecture hall was collected from three preexisting WiFi routers without any modifications or additional infrastructure. The number of unique users connected to the WiFi routers is regarded as the number of occupants present in the hall. Occupant identifiers such as MAC (Media Access Control) addresses and user IDs were filtered out of the data. Detailed analysis of reliability, accuracy and validation processes of the WiFi router data were presented in (Simma et al. 2019, Simma et al. 2019).

Data for this study was collected in the Fall semester of 2019 in the Mechanical Engineering building at the University of New Mexico. The semester consisted of 17 weeks starting from August 19 through December 14. From these 17 weeks, only the weekdays are analyzed as the lecture hall remained unoccupied on the weekends. During the week, Mondays, Wednesdays, and Fridays shared similar class schedules with four lectures that took place throughout the day.

Similarly, Tuesdays and Thursdays shared identical class schedules with five lectures on each day. Therefore, the collected data consisted of 17 occupancy profiles for each weekday.

2.2.Data Normalization

The data for each 24-hour weekday was divided into 200 seven-minute timesteps starting at midnight. The Fuzzy ART neural network requires real valued inputs in the [0,1] interval for stable learning. Therefore, all occupancy data were mapped to values between 0-1. The normalization process was done using the maximum occupancy of the lecture hall (i.e., 100) and was uniformly applied to all 17 weeks of data. A sample of normalized data for one week (Mon-Fri) is shown in Figure 14. The Fuzzy ART input are occupancy profile vectors with 200 points each.

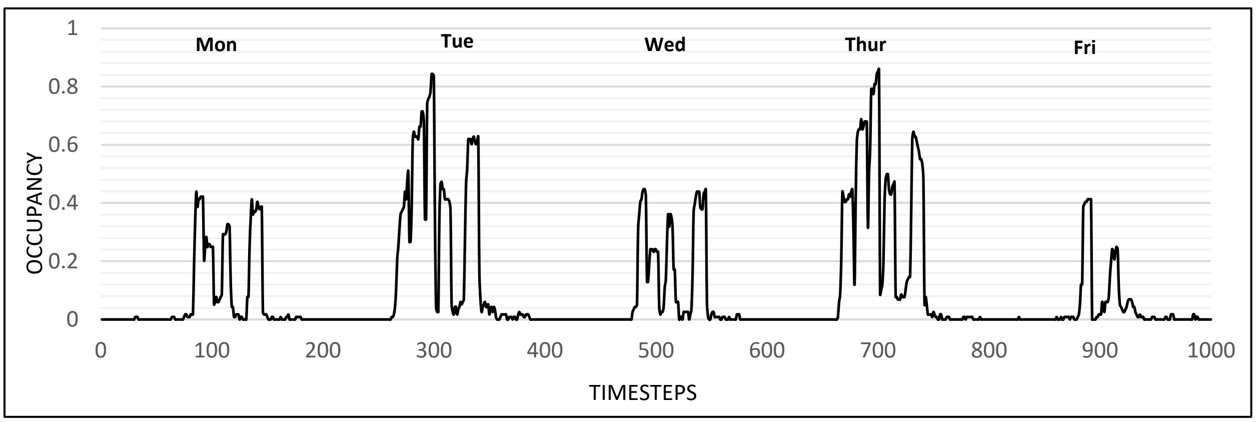


Figure 14: Normalized 5-day Week Occupancy Data

2.3.Fuzzy ART Clusters

Figure 15 shows four consecutive 24-hour Monday occupancy profiles of the lecture hall. From the figure it is evident that even though the scheduled occupancy on Mondays is identical, the occupancy profiles show variations. Similarly, the rest of the Mondays in the dataset showed variations. Fuzzy ART is capable of clustering these profiles into categories and visually represent them by templates using a granularity control parameter called Vigilance (ρ) that ranges between 0 and 1.

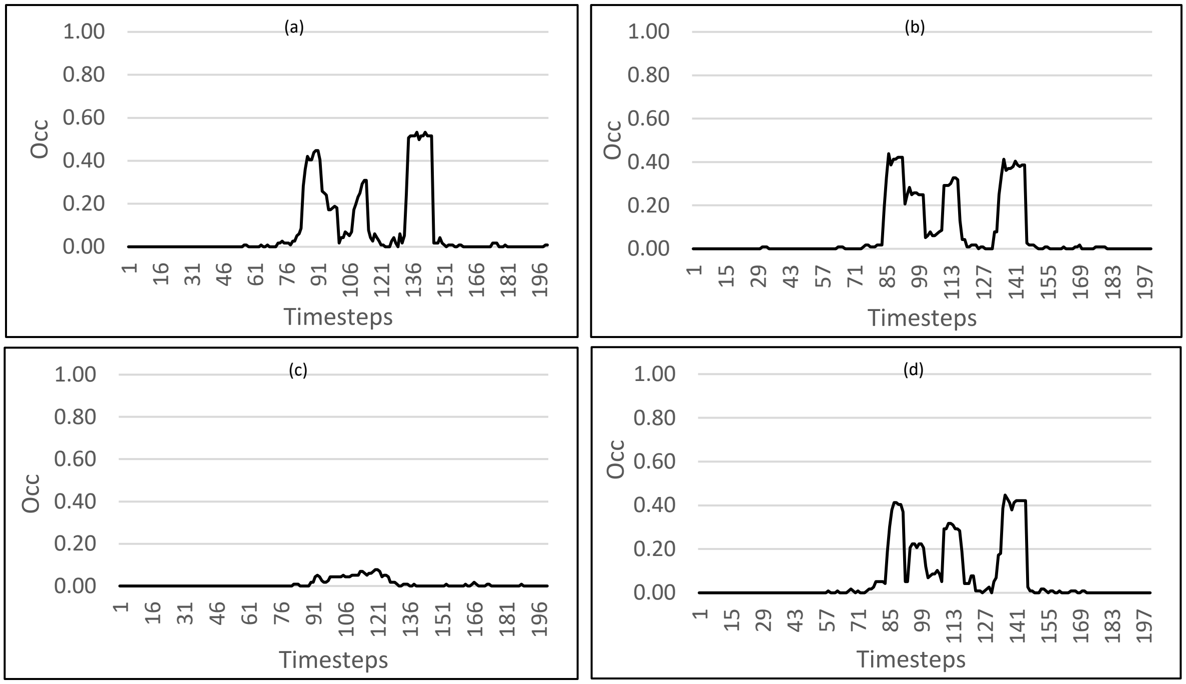


Figure 15: Monday occupancy profiles, a) Aug-19, b) Aug-26, c) Sep-02, d) Sep-09

Vigilance determines the tolerance of variability between occupancy profiles that go into each category. Whenever the variability in the input profiles exceed the set tolerance, a new category is created. As the ρ value increases the tolerance for variability decreases resulting in the creation of more categories. Therefore, at ρ value 1, maximum number of categories are learned, often equal to total samples in the set.

2.4.Output Processing and Pattern Analysis

Given a ρ values, the neural network algorithm clusters similar occupancy profiles represented by a fuzzy template. In this case, a 200x1 daily occupancy input vector (i.e., 24-hour profiles in Figure 15) was presented to the ART algorithm. The algorithm then creates a fuzzy template category that consists of the minimum and maximum bounds for the profiles. The fuzzy template is essentially the envelope of the member profiles of a category. In the case discussed here, a daily template (i.e., a 200-dimensional hyperbox) consists of 200 minimum and maximum values one for each timestep in the profile.

Using the data collected, three different experiments were conducted: 1) Individual weekdays analyzed separately (e.g., 17 samples of 24-hour occupancy input vectors for Mondays are presented to Fuzzy ART), 2) The entire semester data (i.e., 85 samples of 200x1 24-hour occupancy input vectors collected over 17 weeks) analyzed, and 3) samples of time segments defined by lecture start and end time analyzed. A heatmap is created with templates generated. Each template is assigned a unique color to visualize pattern variation across the semester.

3. Results and Discussion

The data collected was normalized as explained in the Methodology and presented to the Fuzzy ART algorithm in various combinations for the experiments listed in Section 2.3. As explained in the methodology, despite identical occupancy schedules for Mondays, Figure 15 shows slightly different patterns of occupancy on four consecutive Mondays. Therefore, for the first experiment, all 17 Monday profiles are presented to Fuzzy ART.

3.1.Experiment-1

Five Fuzzy ARTs are trained separately, one for each day of the week. As explained in Section 2.4, the output consists of templates representing different patterns of occupancy profiles in the data. For example, the four templates learned for 17 Mondays for ρ value 0.92 are shown in **Figure 16**. The four templates learned represent four different occupancy patterns observed among 17 Monday profiles.

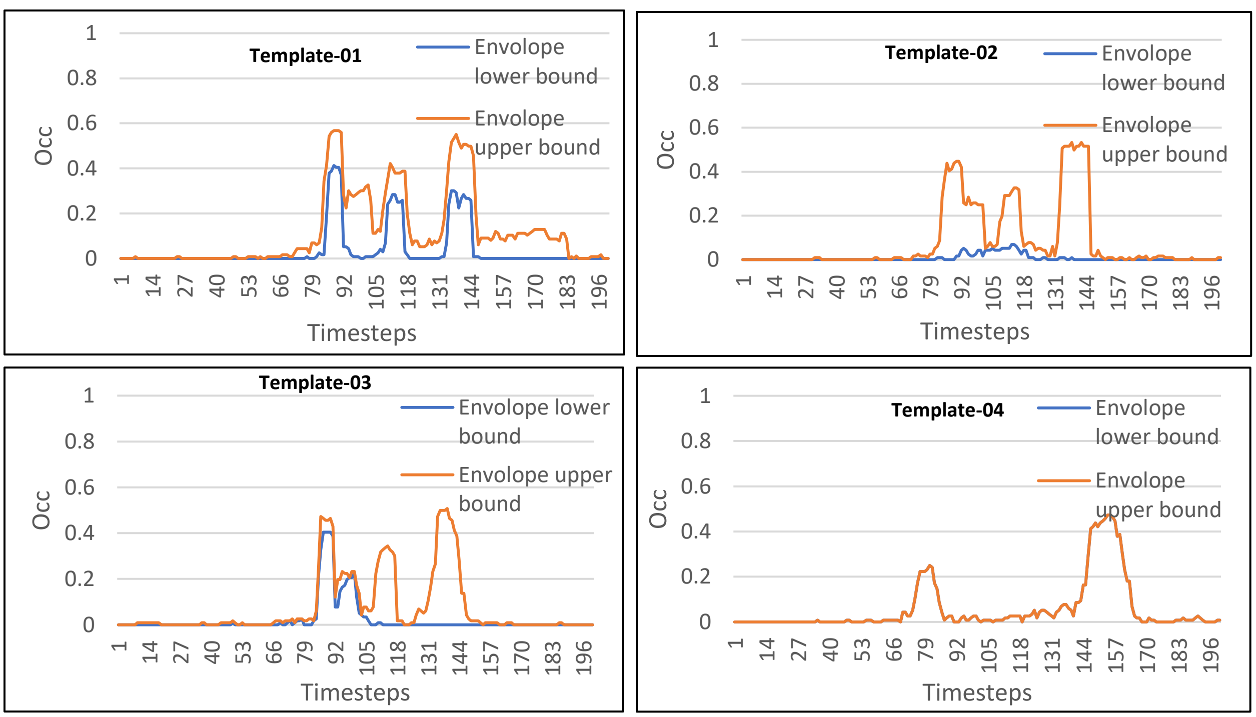


Figure 16: Four templates learned by Fuzzy ART from 17 Mondays input @ ρ = 0.92

However, in Figure 16 template-02, the minimum and maximum curves are widely spread (e.g., between x-axis value 130-148 the minimum is 0 and the maximum is above 0.50). The occupancy predictions are generated using a central curve between the minimum and maximums of the templates. Widely separated bounds of the template indicated a large variation of occupancy around the central curve, implying that the template is too general to be used for prediction. Therefore, the templates learned need to have tighter boundaries (i.e., smaller range between minimum and maximum) for accurate predictions.

Unlike template-01 in Figure 16, template-02 shows significant variability in the profiles that are clustered. These patterns are not ideal to be used in prediction studies as the predicted value might fall over a wide region between minimum and maximum. Additionally, template-04 represents a ‘Point Template (PT)’, meaning only one profile is represented in this category making it a unique occupancy profile. To address the over generalized wide boundary issue, higher ρ values were used during the learning process. As explained in Section 2.3, the number of patterns detected will increase with ρ, making more templates that are less generalized (i.e., tighter bounds on the template enveloped, less variation). For example, ρ values 0.97 and 0.98 generated 9 and 14 different templates respectively from 17 Monday input profiles. Out of the 14 categories learned at ρ value 0.98, 11 are point templates. This infers that the occupancy patterns in the input data are less repetitive.

Similarly, the templates generated for rest of the weekdays at different ρ values revealed similar issues. To visualize the impact of vigilance over number of templates learned (i.e., patterns detected), a heatmap was created with templates learned for each weekday Neural Network (NN). **Table 5** (i) and (ii) show the heatmap for five NNs at ρ values 0.92 and 0.95 where **Table 5** (ii) shows a higher number of templates than in **Table 5** (i). Further increment in ρ value resulted in increased number of templates from each NN. In the case of Tuesdays and Thursdays, 15 and 13 templates were learned respectively from 17 input profiles indicating that the input occupancy profiles have significant variability.

Table 5: Heatmap of templates at ρ = 0.920 (i) and ρ = 0.950 (ii). Each color corresponds to a template.

ρ = 0.92					
	Mon	Tue	Wed	Thur	Fri
wk-1	t1_m	t1_t	t1_w	t1tr	t1_f
wk-2	t1_m	t1_t	t1_w	t1_tr	t1_f
wk-3	t1_m	t1_t	t1_w	t1_tr	t1_f
wk-4	t2_m	t1_t	t1_w	t2_tr	t1_f
wk-5	t2_m	t1_t	t1_w	t2_tr	t1_f
wk-6	t2_m	t2_t	t1_w	t3_tr	t2_f
wk-7	t2_m	t2_t	t1_w	t3_tr	t1_f
wk-8	t2_m	t2_t	t1_w	t4_tr	t2_f
wk-9	t2_m	t2_t	t1_w	t2_tr	t1_f
wk-10	t2_m	t3_t	t1_w	t5_tr	t1_f
wk-11	t2_m	t3_t	t1_w	t5_tr	t3_f
wk-12	t2_m	t4_t	t1_w	t5_tr	t2_f
wk-13	t2_m	t5_t	t1_w	t6_tr	t2_f
wk-14	t3_m	t3_t	t1_w	t6_tr	t2_f
wk-15	t2_m	t3_t	t2_w	t4_tr	t2_f
wk-16	t3_m	t5_t	t1_w	t4_tr	t3_f
wk-17	t4_m	t6_t	t2_w	t4_tr	t4_f

ρ = 0.95					
	Mon	Tue	Wed	Thur	Fri
wk-1	t1_m	t1_t	t1_w	t1_tr	t1_f
wk-2	t1_m	t1_t	t1_w	t1_tr	t1_f
wk-3	t2_m	t2_t	t1_w	t2_tr	t2_f
wk-4	t1_m	t2_t	t1_w	t3_tr	t1_f
wk-5	t1_m	t3_t	t1_w	t4_tr	t2_f
wk-6	t1_m	t3_t	t2_w	t5_tr	t3_f
wk-7	t3_m	t4_t	t2_w	t6_tr	t1_f
wk-8	t3_m	t4_t	t2_w	t7_tr	t4_f
wk-9	t4_m	t5_t	t1_w	t6_tr	t1_f
wk-10	t1_m	t5_t	t1_w	t8_tr	t1_f
wk-11	t4_m	t6_t	t2_w	t4_tr	t3_f
wk-12	t4_m	t7_t	t3_w	t9_tr	t4_f
wk-13	t4_m	t8_t	t3_w	t9_tr	t3_f
wk-14	t5_m	t6_t	t3_w	t8_tr	t5_f
wk-15	t4_m	t6_t	t4_w	t7_tr	t4_f
wk-16	t2_m	t9_t	t3_w	t5_tr	t4_f
wk-17	t6_m	t10_t	t4_w	t7_tr	t6_f

3.2.Experiment-2

The lecture hall has a scheduled occupancy of registered students with ten different classes repeating on multiple weekdays as explained in Section 2.1. Therefore, the occupancy profiles should be similar on days with similar scheduled occupancy. For example, Mondays, Wednesdays, and Fridays share comparable scheduled occupancy. Tuesdays and Thursdays share identical scheduled occupancy. From Experiment-1, it is evident that individual weekdays (e.g., Templates from 17-Monday Fuzzy ART) showed little repetitiveness. Therefore, analyzing the whole data set could identify patterns from days that share similar scheduled occupancy. To that extent, in this experiment occupancy data from all weekdays (i.e., 17 weeks of 5-weekday data) were used in one Fuzzy ART.

Two different input orders were tested in this experiment to ensure that the order of input does not affect the patterns learned. The first order started the input with 17-Mondays followed by 17-Tuesdays, and so on and the second input order had week-1 data (Mon, Tue, Wed, and so on) followed by week-2, and so on up to week-17. The first learning order at ρ values 0.90 and 0.96 generated 14 and 42 categories, respectively. Similarly, the second learning order at ρ values 0.90 and 0.96 generated 15 and 41 categories, respectively. The heatmaps shown in **Table 6** (i) and (ii) are for learning order one at ρ values 0.90 and 0.96. In both learning orders tested, the NN learned over 40 templates out of which at least 14 are point templates. These results indicate that the majority of the 24-hour profiles do not resonate with other profiles.

Table 6: Experiment-2 full dataset heatmaps for ρ = 0.90 (i) and ρ = 0.96 (ii)

ρ = 0.90					
	Mon	Tue	Wed	Thur	Fri
Wk-1	t1	t4	t2	t9	t8
Wk-2	t1	t4	t1	t9	t3
Wk-3	t1	t4	t8	t9	t7
Wk-4	t1	t4	t2	t6	t12
Wk-5	t1	t4	t2	t9	t12
Wk-6	t1	t4	t8	t6	t12
Wk-7	t1	t4	t8	t6	t12
Wk-8	t2	t5	t8	t3	t3
Wk-9	t1	t4	t8	t10	t12
Wk-10	t1	t5	t2	t10	t1
Wk-11	t2	t5	t2	t5	t12
Wk-12	t2	t6	t2	t11	t13
Wk-13	t1	t5	t8	t11	t12
Wk-14	t2	t6	t8	t11	t12
Wk-15	t2	t5	t8	t3	t3
Wk-16	t2	t7	t8	t12	t8
Wk-17	t3	t7	t3	t3	t14

ρ = 0.96					
	Mon	Tue	Wed	Thur	Fri
Wk-1	t1	t8	t6	t25	t23
Wk-2	t1	t8	t6	t25	t36
Wk-3	t2	t9	t19	t26	t36
Wk-4	t1	t9	t6	t27	t37
Wk-5	t3	t10	t6	t28	t37
Wk-6	t3	t11	t19	t29	t38
Wk-7	t3	t10	t20	t30	t37
Wk-8	t4	t12	t20	t2	t2
Wk-9	t4	t12	t21	t31	t37
Wk-10	t1	t13	t21	t32	t38
Wk-11	t5	t14	t21	t33	t39
Wk-12	t5	t15	t3	t34	t40
Wk-13	t5	t16	t22	t35	t40
Wk-14	t6	t14	t22	t31	t40
Wk-15	t5	t13	t23	t2	t2
Wk-16	t2	t17	t22	t29	t41
Wk-17	t7	t18	t24	t24	t42

3.3.Experiment-3

Figure 17 shows a template with a section of the occupancy profile with a wide interval. The reasons for such wide intervals during a specific part of the templates are attributed to unscheduled class cancellations, class relocation, and other interruptions. This is an indication that each class period has its own statistics. The two prior occupied segments in Figure 17 with tighter minimum and maximum intervals confirming this hypothesis. As explained earlier in Section 3.1, templates with tighter bounds are required for accurate prediction. Since parts of the 24-hour profiles have their own statistics, the full day profiles are subdivided into individual class segments (aka time-segments). A time segment defines the interval of time in which the lecture hall is scheduled to be occupied. The lecture hall has ten different classes scheduled in a week with different occupancy volumes resulting in ten occupied time segments over the week.

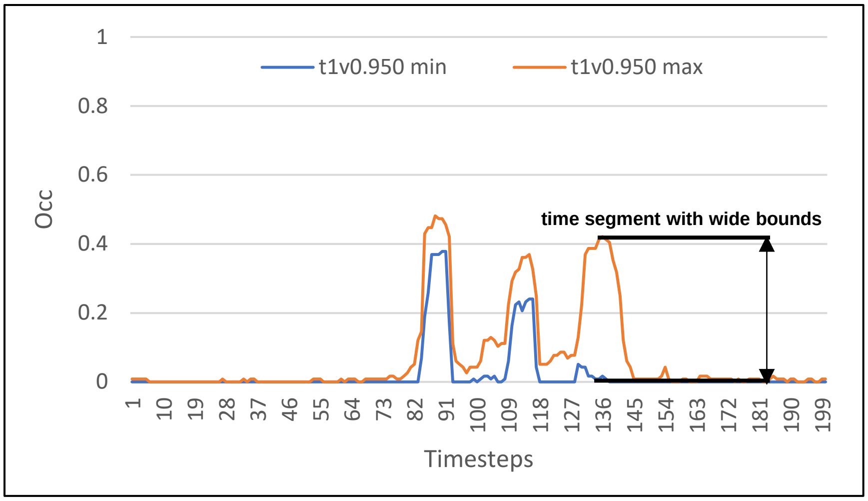


Figure 17: 24-Hour templates with wide Min and Max interval

In this experiment, individual time segments for each class were presented to Fuzzy ART. Additionally, the 17th week was removed from the analysis as it was designated as finals week having statistical variability independent of the previous 16 regular weeks. Templates learned for time segment 1 (i.e., class-01 scheduled on MWF from 10:00am – 10:50am) at ρ value 0.80 are shown in Figure 18.

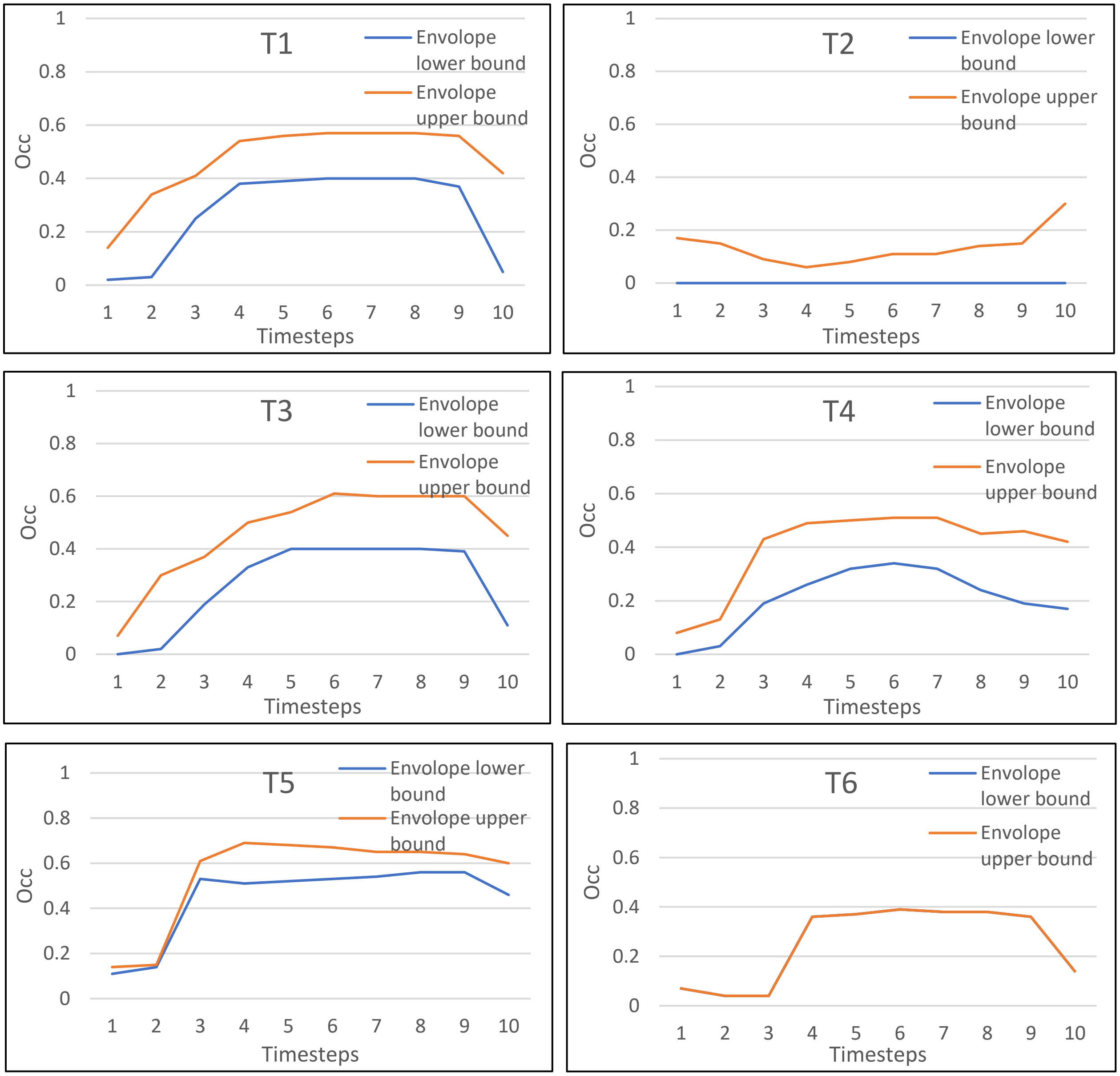


Figure 18: Six templates learned from time segment - 01 (i.e., class-01)

The templates learned for individual time segments have tighter bounds unlike the templates learned in previous experiments. The heatmap shown in **Table 7** demonstrates that time intervals with similar occupancy schedule do follow a pattern. From the heatmap, Template-1 can be observed at least once on Mondays, Wednesdays, and Fridays. Additionally, the holiday/cancelled class template (i.e., Template-2) can be visualized from the heatmap. This implies that the patterns detected are less generalized and can be reliably used in occupancy forecast studies.

Table 7: Heatmap for time segment-01. Each color corresponds to a template.

	ρ = 0.80		
	Mon	Wed	Fri
wk-1	t1	t1	t3
wk-2	t1	t1	t3
wk-3	t2	t3	t2
wk-4	t1	t3	t3
wk-5	t1	t3	t5
wk-6	t1	t4	t4
wk-7	t1	t1	t5
wk-8	t1	t5	t2
wk-9	t1	t4	t1
wk-10	t3	t1	t3
wk-11	t3	t1	t6
wk-12	t1	t1	t2
wk-13	t3	t3	t5
wk-14	t3	t3	t1
wk-15	t3	t2	t2
wk-16	t3	t4	t2

4. Conclusion

As hypothesized, Fuzzy ART was able to detect different patterns in the occupancy profiles with similar scheduled occupancy. The experiments further demonstrated that 24-hour occupancy profiles may consists of segments that are statistically independent. If these 24-hour profiles can be subdivided into smaller time segments defined by external schedules or constraints, significant repeating occupancy patterns can be detected. Additionally, templates learned from the smaller time segments are defined by tighter bounds inferring less generality. Furthermore, the heatmaps developed in this study aided in visualization of pattern variations over time. The current methodology successfully demonstrates that significant occupancy patterns can be learned from WiFi router data collected in complex occupancy environment. As commercial spaces tend to have rush hours with increased occupant volumes during specific part of the day, it is possible to isolate these into individual time segments for pattern detection. These patterns could aid studies that focus on occupancy load forecasts to optimize HVAC operation and improve occupant comfort.

ACKNOWLEDGEMENT

The study acknowledges Dr. C Birk Jones from Sandia National Laboratories, Physical Plant Department, and the IT department of the University of New Mexico.

REFERENCES

Basu, Chandrayee, Girish Ghatikar, and Prateek Bansal. 2014. “Enabling Efficient, Responsive, and Resilient Buildings: A Collaboration Between the United States and India,” no. March.

Carpenter, Gail A., Stephen Grossberg, and David B. Rosen. 1991. “Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System.” *Neural Networks* 4 (6): 759–71. https://doi.org/10.1016/0893-6080(91)90056-B.

Chen, Zhenghua, Chaoyang Jiang, and Lihua Xie. 2018. “Building Occupancy Estimation and Detection: A Review.” *Energy and Buildings* 169: 260–70. https://doi.org/10.1016/j.enbuild.2018.03.084.

Ding, Yan, Wanyue Chen, Shen Wei, and Fan Yang. 2021. “An Occupancy Prediction Model for Campus Buildings Based on the Diversity of Occupancy Patterns.” *Sustainable Cities and Society* 64 (October 2020). https://doi.org/10.1016/j.scs.2020.102533.

Dong, Bing, Burton Andrews, Khee Poh Lam, Michael Höynck, Rui Zhang, Yun Shang Chiou, and Diego Benitez. 2010. "An Information Technology Enabled Sustainability Test-Bed (ITEST) for Occupancy Detection through an Environmental Sensing Network." *Energy and Buildings* 42 (7): 1038–46. https://doi.org/10.1016/j.enbuild.2010.01.016.

EIA, and Behjat Hojjati. 2019. "Global Energy Consumption Driven by More Electricity in Residential, Commercial Buildings." *U.S. Energy Information Association* 2019: 2019–20. https://www.eia.gov/todayinenergy/detail.php?id=41753.

Gaetani, Isabella, Pieter Jan Hoes, and Jan L.M. Hensen. 2016. "Occupant Behavior in Building Energy Simulation: Towards a Fit-for-Purpose Modeling Strategy." *Energy and Buildings* 121: 188–204. https://doi.org/10.1016/j.enbuild.2016.03.038.

Huo, Tengfei, Hong Ren, Xiaoling Zhang, Weiguang Cai, Wei Feng, Nan Zhou, and Xia Wang. 2020. "China ' s Energy Consumption in the Building Sector : A Statistical Yearbook-Energy Balance Sheet Based Splitting Method." *Journal of Cleaner Production* 185 (2018): 665–79. https://doi.org/10.1016/j.jclepro.2018.02.283.

Jiang, Yi, Da Yan, Siyue Guo, and Shan Hu. 2019. "China Building Energy Use 2018," no. December 2016.

Liao, Chenda, Yashen Lin, and Prabir Barooah. 2012. "Agent-Based and Graphical Modelling of Building Occupancy." *Journal of Building Performance Simulation* 5 (1): 5–25. https://doi.org/10.1080/19401493.2010.531143.

Outlook, International Energy. 2017. "Today in Energy Buildings Energy Consumption in India Is Expected to Increase Faster than in Other Regions," 2015–17.

Parys, Wout, Dirk Saelens, and Hugo Hens. 2011. “Coupling of Dynamic Building Simulation with Stochastic Modelling of Occupant Behaviour in Offices - a Review-Based Integrated Methodology.” *Journal of Building Performance Simulation* 4 (4): 339–58. https://doi.org/10.1080/19401493.2010.524711.

Simma, Krishna Chaitanya J, Susan M Bogus, and Andrea Mammoli. 2019. “WI-FI ROUTER NETWORK-BASED OCCUPANCY ESTIMATION TO OPTIMIZE HVAC ENERGY CONSUMPTION” 2015 (Saic 2011): 1–10.

Simma, Krishna Chaitanya Jagadeesh, Andrea Mammoli, and Susan M Bogus. 2019. “Real Time Occupancy Estimation Using WiFi Network to Optimize HVAC Operation” 00 (2018).

U.S. Energy Information Institution. 2018. “Annual Energy Outlook 2018 with Projections to 2050.” *Annual Energy Outlook 2018 with Projections to 2050* 44 (8): 1–64. https://doi.org/DOE/EIA-0383(2012) U.S.

Wang, Wei, Jiayu Chen, and Tianzhen Hong. 2018. “Occupancy Prediction through Machine Learning and Data Fusion of Environmental Sensing and Wi-Fi Sensing in Buildings.” *Automation in Construction* 94 (July): 233–43. https://doi.org/10.1016/j.autcon.2018.07.007.

Chapter-05 – Neural Network Based Occupancy Prediction using Patterns detected in WiFi Occupancy

Krishna Chaitanya Jagadeesh Simma[1], Thomas P Caudell[2], Susan M Bogus[3], Andrea Mammoli[4]

[1]PhD candidate, Department of Civil, Construction and Environmental Engineering, University of New Mexico, MSC01 1070, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: jagadeesh145@unm.edu

[2]Emeritus Professor, Department of Electrical and Computer Engineering, University of New Mexico, MSC01 1070, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: tcaduell@unm.edu

[3]Professor, Department of Civil, Construction and Environmental Engineering, University of New Mexico, MSC01 1070, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: sbogus@unm.edu

[4]Customer Technologies Program, Electric Power Research Institute; 3420 Hillview Ave, Palo Alto, CA 94304; e-mail: amammoli@epri.com

Abstract

The role of occupants in building energy management is well established. Occupant information from commercial buildings can function as a metric for the heating, ventilation, and air-conditioning (HVAC) loads. Predicting these loads ahead in time could aid in setting up HVAC protocols to minimize wasteful energy demand and improve occupant comfort. To this extent, several frameworks were proposed that use occupancy data collected from commercial office buildings and graduate student offices to predict occupancy. However, the occupancy data collected in these frameworks are controlled and have limited complexity compared to larger

commercial spaces such as airports, libraries, lecture halls, and shopping malls. Therefore, this paper presents the results of a study that demonstrates the feasibility of occupancy predictions from a relative dynamic environment using WiFi data and neural networks. The results demonstrate that reliable occupancy predictions are possible using intrinsic variables that have significant correlation with the occupancy profile of a time-segment. Similar conclusion scan be extended to extrinsic variables if they are significantly correlated to the occupancy.

Keywords

Occupancy Prediction, Neural Networks, HVAC, Commercial Buildings, Energy, LAPART

1. Introduction

Integration of technology into various industrial fields is taking place rapidly. The construction industry and built environment (residential and commercial buildings) have seen their share of technology integration in the name of modernization. The investment into modernization of infrastructure recommended in the ASCE Infrastructure Report card[1], led to the surge of renovations, sensor networks, smart devices to control indoor climate, and WiFi infrastructure in most commercial buildings. With buildings consuming over 20% of total delivered energy globally in 2018 [2], the need for energy efficiency through modernization is a priority. In addition, the global commercial building footprint is projected to increase along with the energy consumption per unit area for the next three decades [2][3][4]. In this context, the modernization of commercial buildings with various technologies provides opportunities to achieve an unrealized energy saving in commercial buildings.

The growth in commercial buildings garnered researchers' interest in their energy consumption and their energy-saving potential[5]. This raised the significance of building energy modeling to identify energy efficient designs. However, the actual energy consumption revealed large discrepancies with modeled energy and the reasons for it were attributed to variable building occupancy [6]. Since this finding, commercial building occupancy became an integral part of energy management strategies. The dynamic nature of occupancy within buildings complicates the modeling process [7]. Modernization of buildings with technology aided studies to obtain occupant information from various sensory infrastructure (e.g., environmental sensors, dedicated occupancy sensors, RFID, Bluetooth devices, and WiFi networks) [8]. However, accuracy, complexity, costs associated with dedicated infrastructure and privacy concerns impeded the wide scale implementation of these frameworks.

In commercial buildings, heating, ventilation, and air conditioning (HVAC) accounts for 40% of total energy consumption [9]. A demand-driven HVAC operational strategy can play a significant role in the energy efficiency of commercial buildings [7]. From the results in [10], it is evident that existing WiFi routers can be a source of reliable occupancy estimation without added infrastructure and complex computational resources while preserving occupants' privacy. Additionally, the occupancy-based HVAC schedules and the corresponding energy consumption results published in [11] emphasize that demand-driven HVAC operation can significantly reduce building energy consumption. However, knowing the occupancy profiles of a building in advance can improve energy efficiency furthermore by reducing unnecessary energy demands of the HVAC systems [5],[9]. Additionally, the knowledge of occupant load in advance can be used to pre-condition (heat/cool) the space which can improve occupant thermal comfort levels.

To acquire knowledge of occupancy, prediction frameworks were proposed that frequently employed Markov chains, support vector machine, k-nearest neighbor (kNN), and Artificial Neural Networks (ANN) [12]. Using the data collected from a variety of sensory infrastructure within commercial spaces such as commercial offices, university offices, single, and multi-person offices, occupancy predictions were made at different levels. Some frameworks predicted the state of the occupant ('in' or 'out') [13], others predicted the number of occupants by detecting their state [14],[15], and the rest focused on behavioral aspects (choice of lighting, windows, and blinds) of occupants [16], [17]. While these frameworks provide insight into predicting fixed number of "long-term" occupants in office spaces, occupancy prediction in a dynamic environment (e.g., airports, shopping malls, gymnasiums) remains a challenge. This paper presents the results of a study that demonstrates the feasibility of making short term occupancy prediction in a relatively dynamic environment using WiFi data and neural networks.

2. Literature

Occupant behavior is considered one of the most complex processes taking place within buildings [18]. The stochastic nature of building occupants complicates the process of indoor microclimate control (heating and cooling) and energy conservation strategies. Researchers attempted to model and predict this complex behavior of occupants to achieve various energy efficiency goals over the past decade. Studies used occupants' behavioral aspects such as windows, blinds, and lighting choices to improve energy efficiency [16]. The comprehensive review conducted by [17] highlights that building occupants tend to override any automated window shading, and lighting protocols to suit their personal preference. However, generalizing such behavior of occupants can be tedious and may not always aid in realizing the maximum possible energy savings in commercial buildings. More importantly, in commercial spaces such as airports and university buildings, occupants may not have control over lighting and shading.

Discrete methods as in case of [15] used ‘occupancy matrix’ to define spatial distribution of occupancy in an office space. The occupancy matrix is formed by dividing a 2-D space into multiple zones based on the thermal zones and presence of an occupant in each zone. A binary system of zeros and ones was used to represent occupancy and non-occupancy zones. Occupancy data from the test space was obtained using dedicated iBeacon infrastructure. This occupancy data coupled with the thermal zone information defined the dynamic spatial occupancy distribution. Using the spatial occupancy distribution (i.e., occupancy matrix) cooling requirements for unoccupied zones was reduced. However, office spaces are typically occupied by same set of occupants during the occupied state and this approach may not be applicable for larger commercial spaces with new occupants every day.

Environmental sensors such as CO_2 sensors use the correlation between the standard quantity (e.g., CO_2/person) per person to estimate occupancy. In [19], occupancy data from CO_2 sensors were used with hidden Markov model algorithm to predict occupancy. This system utilizes the correlation between occupants and CO_2 concentrations to predict occupancy. Using this approach, the prediction made for an office space achieved an accuracy ranging between 85% - 93%. While the occupancy inferred is non-intrusive in nature, environmental sensors such as CO_2 sensors can be unreliable in occupancy estimation due to the complexity of CO_2 concentrations which could results in inaccurate predictions [5]. Additionally, the results are from tests performed for a maximum of five occupants and the prediction accuracy may differ in dynamic environment with higher number of occupants.

In a different approach, WiFi probe data (i.e., media access control (MAC) address) was used in tandem with Markov based feedback recurrent neural network algorithm to predict occupancy [5]. A graduate student office with 25 ‘long-term’ residents was used to conduct various validation experiments. Each MAC address was assigned to a single occupant and that information was stored as memory. Using the neural network approach, the state of each MAC address ("in" or "out") was predicted and thereby estimate the total occupancy of the office space. Contingent on allowed tolerances, this approached reached accuracy ranging between 80% to 94%. However, in a dynamic occupancy environment (e.g., airports and shopping malls) a MAC address assignment approach is not possible as occupants of such commercial spaces cannot be classified as 'long-term' residents. Additionally, MAC addresses contain occupant identifiers and may raise occupant privacy concerns.

A separate set of studies proposed frameworks for using occupancy data obtained from different sources such as RFID tags [20], cameras [8], and smart meters [21], among others to improve energy efficiency of commercial buildings. In summary, these frameworks achieve their goals with limitations such as privacy concerns, added infrastructure expenses, and limited applicability, among others. These limitations may hinder large scale implementation of these frameworks and when implemented may not realize the full potential in energy savings.

Occupant comfort level data highlight the need for HVAC energy conserving strategies to incorporate occupant comfort parameters [22]. The heat dissipation of occupants inside buildings contribute towards the cooling load in the cooling dominated months and [23][15]. This occupant heat dissipation effect has a profound impact on heating loads in heating dominated months [23]. Additionally, the metabolic rates of occupants vary depending on the type of commercial building. The metabolic rates of occupants in gymnasiums, stores, and terminal buildings are reported to be higher at 1.6met (metabolic equivalent unit) followed by schools at 1.2met compared to other types of commercial buildings [24]. The higher metabolic rates induce higher heat loss of occupants due to their homotherm nature [25]. The metabolic rates and the amount of heat dissipated by occupants in both summer and winter months are identical [24], highlighting that occupants' thermal interaction in all types of weathers remain constant. Furthermore, the human body radiates different levels of heat throughout the day, implying that the occupants' thermal interactions within buildings are dynamic throughout the day [26]. Moreover, dynamic variation of indoor spatial temperature that account for thermo-physiological parameters such as occupants' metabolic rate and sweat production can positively impact their comfort and health [25].

In summary, a constant heating or cooling setpoint temperatures may not improve occupants' thermal comforts inside buildings. A new setpoint that is a function of number of occupants in a space may contribute towards improving occupant comfort inside buildings. Additionally, knowing the occupant count for a space ahead in time can allow for pre-conditioning the space to that specific number of people can result in increased occupant comfort and avoid HVAC wasteful demand. With emphasis on the need for such HVAC strategies [27], [28], an occupancy prediction approach that addresses the identified limitations is essential.

To this extent, predicting occupancy from a dynamic environment such as airports, public libraries, institutional buildings, and gymnasiums among others can be used to create HVAC operational schedules to improve occupants' thermal comfort and reduce energy consumption. In this study, using patterns detected in WiFi measured occupancy [29] and the LAPART neural network [30], occupancy predictions were made for a university lecture hall that has relatively complex occupancy patterns.

3. Methodology

The methodology for this study consists of three parts: 1) Occupancy pattern detection: Using a Fuzzy ART Neural Network to learn patterns in WiFi measured occupancy data, 2) Occupancy prediction: Using a LAPART neural network to make short term occupancy profile predictions, and 3) Energy Calculations: HVAC energy consumption was estimated using predicted occupancy profiles. **Figure 19** graphically illustrates the components of the methodology.

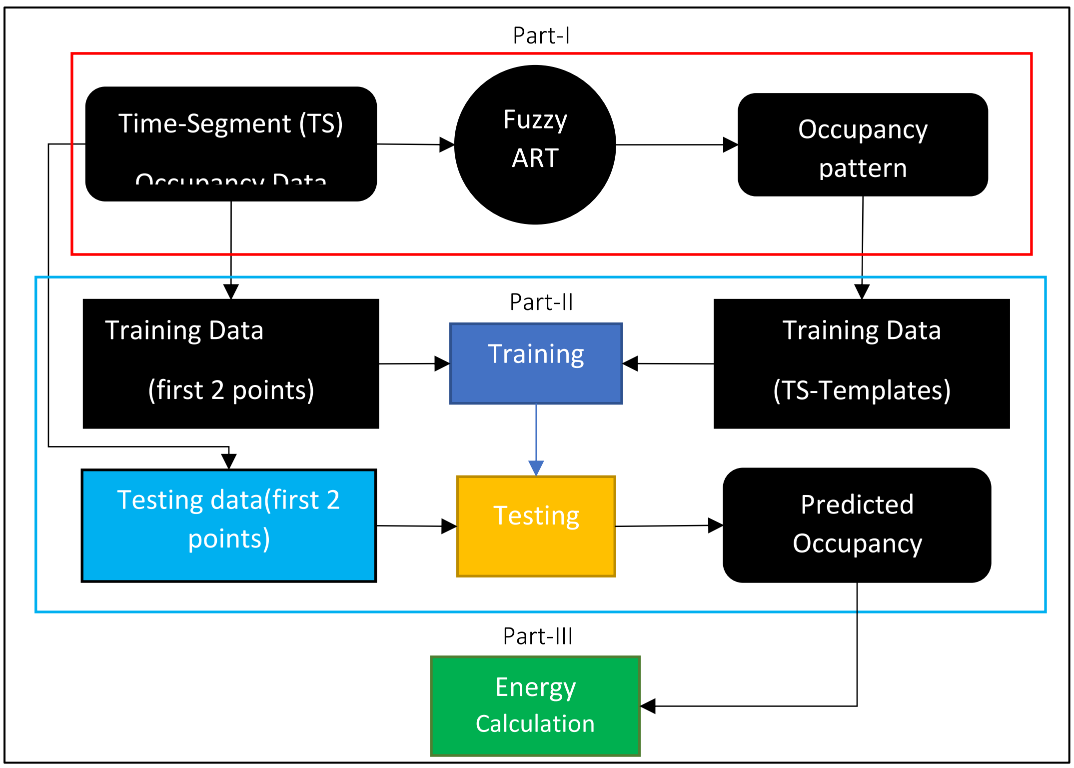


Figure 19: Methodology

3.1.Part-I: Occupancy pattern detection

Sixteen weeks of occupancy data were collected from existing WiFi routers in a lecture hall at University of New Mexico between August 18, 2019, and December 07, 2019. The structure of patterns in the WiFi measured occupancy need to be studied to learn the statistical variability in the data. This allows for learning rules by which occupancy profiles represented by input patterns can be classified into various categories [31]. Patterns in the occupancy data were learned using the Fuzzy ART Neural Network [32]. This self-organizing neural network segments data into unlabeled categories, the granularity (i.e., size of the hyperbox) of which is determined by a single vigilance parameter, ρ. The effect of ρ on number of categories learned is visualized in **Figure 20**.

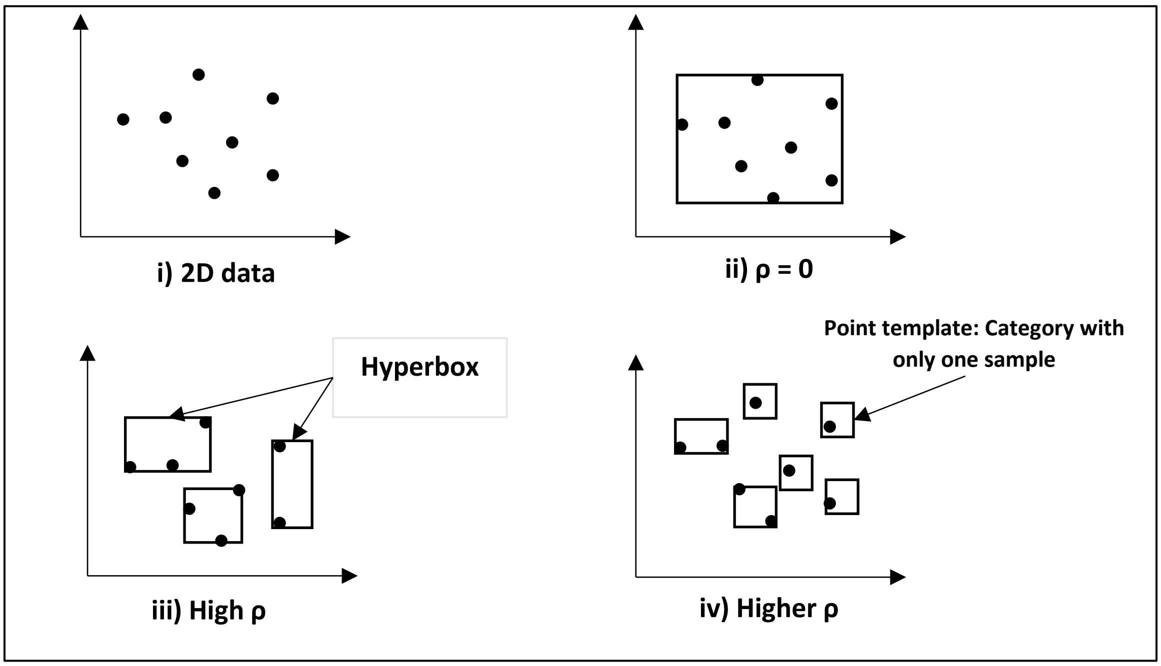
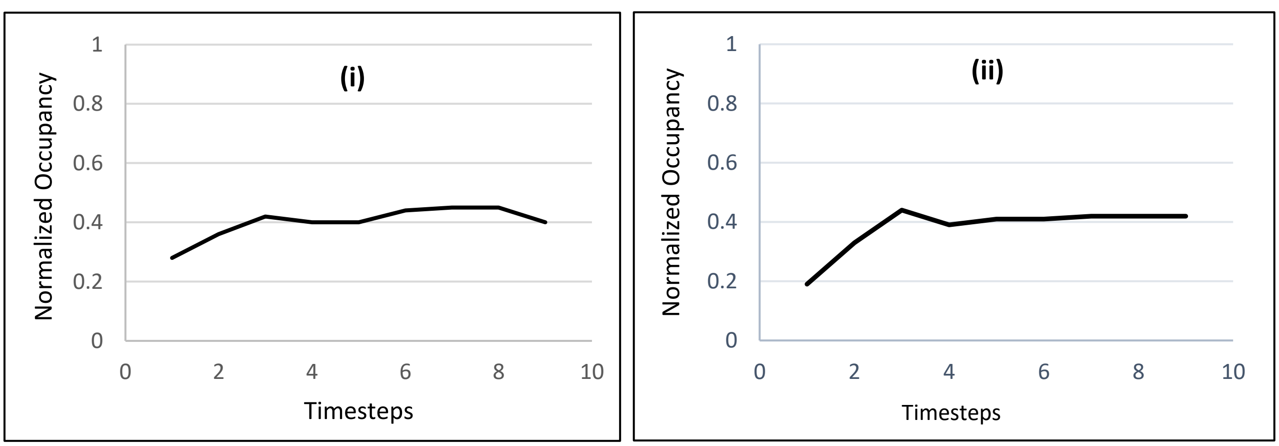


Figure 20: Fuzzy ART, effect of ρ on hyperboxes. i) synthetic 2D data, ii) hyperbox at ρ = 0, iii) hyperboxes at a high ρ (e.g., 0.8), and iv) hyperboxes at a higher ρ value (e.g., 0.95)

In this study, each category represents a collection of 'similar' occupancy profiles. Fuzzy ART executes an unsupervised learning process and for every input profile (examples shown in **Figure 21**, **i) and ii**), the neural network searches for a category that it matches [33]. **Figure 21**. Plots are in a 9D space where, the timestep axis represents the orthogonal axes of 2D space plots shown in figure 2. The learned categories are visually represented by a fuzzy template shown in **Figure 21**, **(iii)** (i.e., n-dimensional hyperbox) and the template selected by the neural network that best matches the input profiles is referred to as the matching template. The fuzzy templates are defined by the minimum and maximum of member profiles. Essentially, each fuzzy template is an envelope of member profiles.

Various experiments were performed to learn patterns in the occupancy data and the results were presented in [34]. The results in [34] emphasize that significant patterns can be learned from WiFi measured occupancy data by dividing a 24-hour profile into smaller time-segments bound by external schedules. The time-segments in this study were defined by the duration of lectures scheduled. The scheduled lectures repeat twice or thrice in a week. For examples, lectures that are scheduled for 50 minutes repeat three times in a week and lectures that are scheduled for 75 minutes repeat twice in a week. The occupancy profiles for 50-minute and 75-minute time-segments are represented by a 9-D and 12-D vectors, respectively. Ten different time-segments were identified throughout the data collected. Occupancy profiles of individual time-segments are clustered into various categories by the Fuzzy ART neural network algorithm. The hyperbox learned to enclose the data points of a category will be referred to as a template for the remainder of this paper.



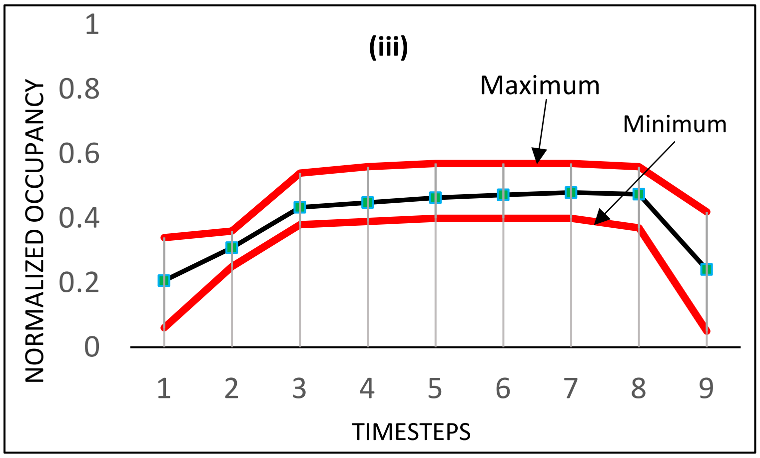


Figure 21: i) sample input profile-1, ii) sample input profile-2, iii) A sample fuzzy template representing category of similar profiles.

From the templates learned for each time-segment, the one that occurred most frequently was identified. The fuzzy template representing this frequently repeating template is termed as *Most Likely Template* (MLT). The time-segments identified in the data take place on multiple days. For example, time-segment #1 is a 50-minute time-segment that was scheduled on Mondays, Wednesdays, and Fridays from 10:00 am to 10:50am. The MLT for a time-segment can be different on each scheduled day (i.e., for time-segment #1, the MLT on Mondays and Wednesdays is Template #1 but for Fridays, it is Template #2). Similarly, the MLTs for each time-segment were identified on all scheduled days. Additionally, each time-segment has a template with zero occupancy profiles (i.e., occupancy remained zero throughout the time-segment) and it is termed as *Holiday Template* (HT). HT profiles represent holidays, lecture cancellations, and relocations, among others.

Furthermore, HVAC energy consumption for a 5-day week was estimated for four different occupancy schedules: 1) MLT Schedule: a composite profile for each day of the week was created by concatenating the MLT average profiles (i.e., average of measured member occupancy profiles) for all the time-segments in a day, 2) WiFi Schedule: occupancy profile with actual WiFi measured samples for each time-segment, 3) Registered Schedule: occupancy profile with registered number of students for each lecture (i.e., time-segments) scheduled in a day, and 4) Fixed Schedule: a fixed schedule that assumes maximum occupancy during working hours (from the beginning of the first time-segment to the end of last time-segment in a day). These energy simulations were performed in EnergyPlus using a dual-duct energy model. Detailed steps for dual-duct energy simulation setup are explained in Section 3.3. The HVAC energy consumption results can corroborate the potential use of MLTs as expected occupancy for a time-segment.

3.2. Part-II: Occupancy prediction

The MLTs identified for each time-segment may be used as expected occupancy. As explained in section 3.1, multiple templates were learned for each time-segment as shown in **Figure 22**. Each profile shown in **Figure 22** gives an example of the average occupancy profiles (i.e., average of measured member occupancy profiles) of the six templates learned for time-segment #4. Out of the six occupancy profiles shown in **Figure 22**, Template #1 (shown in blue dashed line) was identified as the time-segment's MLT, and Template #2 (shown in red) was identified as the HT. Hypothetically, when MLT profile is used as expected occupancy for time-segment #4 and the actual occupancy matches the HT profile, the occupancy is overestimated.

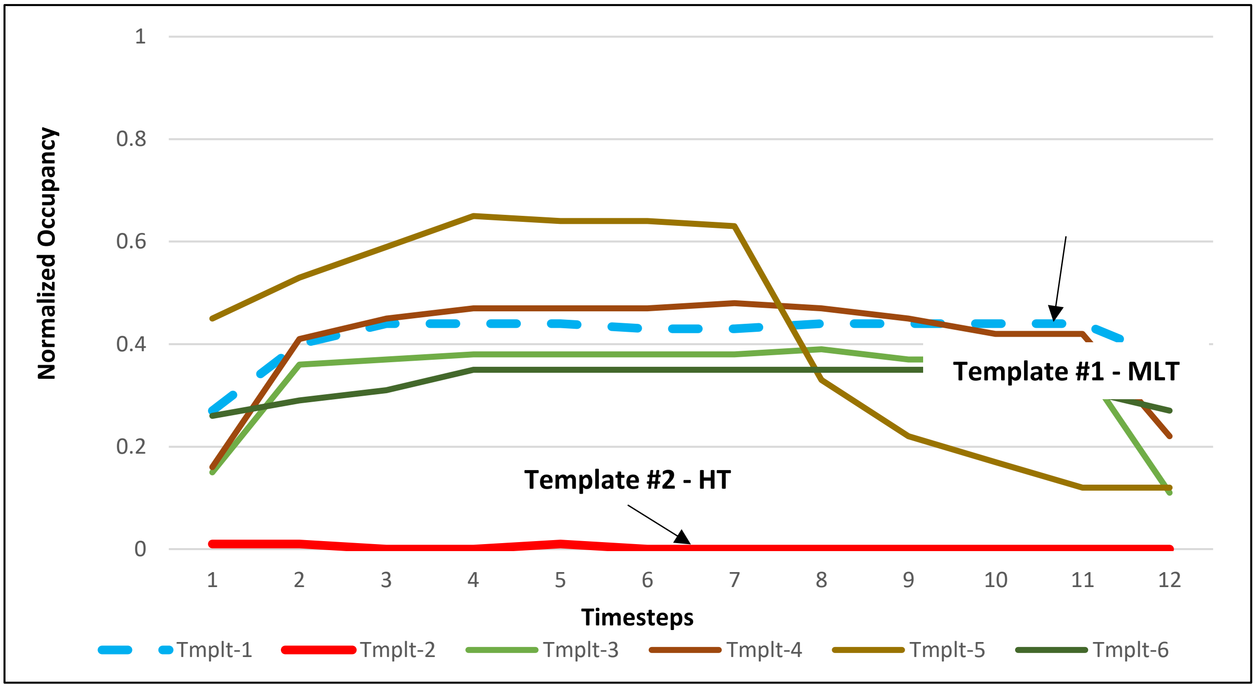


Figure 22: Learned template profiles (average profiles) for time-segment #4.

Additionally, for time-segment #1, the MLT on Fridays is Template #2 as explained in section 3.1. However, Template #2 is also the HT for time-segment #1. Therefore, when Template #2 profile is used as expected occupancy, the occupancy maybe underestimated relative to the actual occupancy. **Figure 23** shows the MLT and HT profiles along with the rest of the learned profiles for time-segment#1.

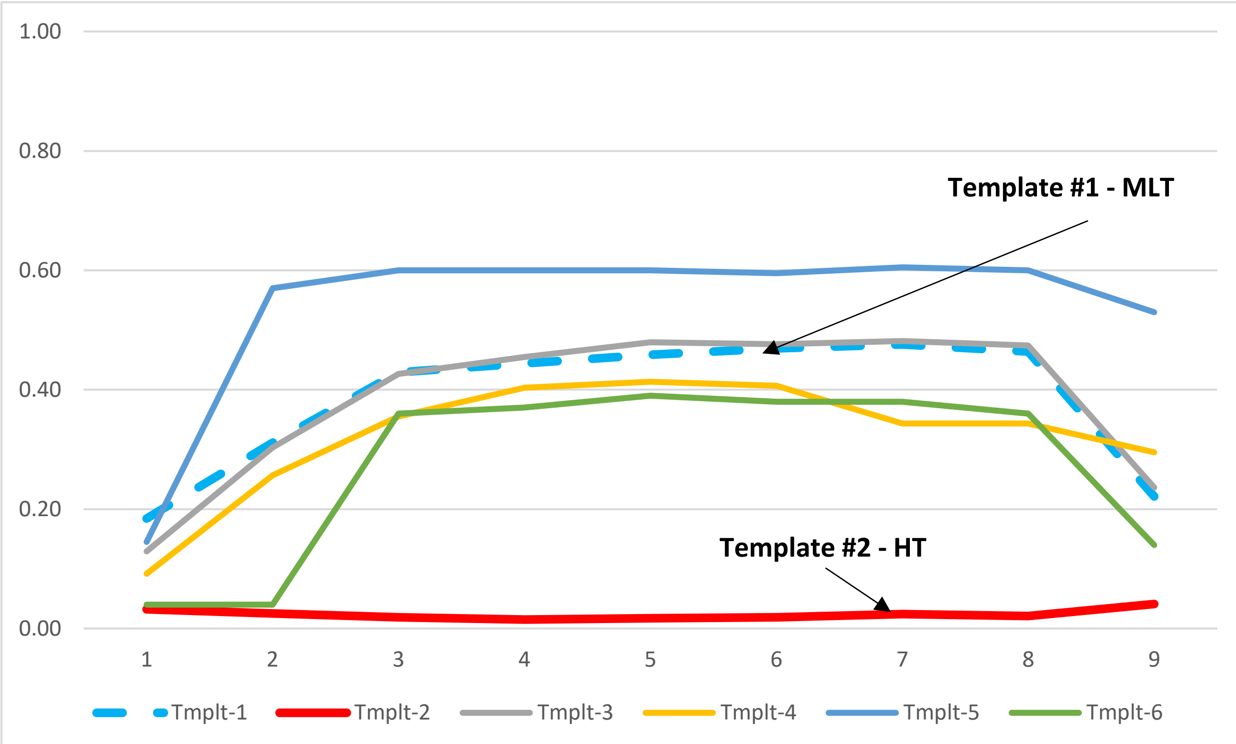


Figure 23: Template Profiles (average profiles) for Time-Segment #1

Therefore, operating HVAC systems based on a fixed MLT profile as expected occupancy may result in wasteful HVAC demand when occupancy is overestimated and occupants' discomfort when occupancy is underestimated. A mechanism is needed to able to detect a better suited template for a time-segment when MLT is incorrect. Intrinsic and extrinsic parameters available can be used to detect a more suited template profile. Some of the extrinsic variables that can be used are day of week, time of the day, occupancy schedules, and other parameters that are unique to a specific type of commercial building. Extrinsic variables such as day of the week and occupancy schedules were used to select the MLT profiles for a time-segment. Intrinsic variables may hold sufficient information to be able to find a better matching template. For this study, the first two samples of real-time WiFi measured occupancy at the start of each time-segment may be used to find a better matching template for the reminder of the time-segment. Therefore, the feasibility of using the first two real-time WiFi measured occupancy need to be verified. To this extent, using the first two time-samples of a time-segment, the 2D Euclidean distance between the first two points of its resonant template and the MLT were compared with the Euclidean distance of the full resonant template with the MLT distance. If these distances are highly correlated then, it is possible that the initial time-segment will be predictive of the full-time segment. The data used to test the correlations were calculated using **Equation 1**.

Equation 1: Equation to calculate the multidimensional distance. For a 2-D point (n = 2); 9-D point (n = 9); 12-D point (n = 12); x_i = n-dimensional MLT profile; a_i = n-dimensional template profile

$$\text{Multidimension distance per template w.r.t MLT} = \sqrt{\sum_{i=1}^{n}(x_i - a_i)^2}$$

For 50-min time-segments, distances in 2D and 9D spaces were computed and for 75-min time-segments, distances in 2D and 12D spaces were computed. A total of *n-1* pairs of distances were calculated for each time-segment where *n* is the number of templates learned for that time-segment. The calculated points were plotted against each other and the R^2 values were calculated. The R^2 value represents the part of the variance of a dependent variable (i.e., full template) explained by an independent variable (i.e., the first two template samples). Here a high R^2 value implies the possibility of distinguishing the templates using the first two template time-samples. Therefore, a high correlation between the distances calculated in different spaces lends evidence that it may be possible to learn to associate the 2D points with 9D points. Based on this hypothesis, using Neural Networks and the first two measured occupancy-samples of a time-segment a matching template was predicted for that time-segment. The template predictions were made using Laterally Primed Adaptive Resonance Theory (LAPART) neural network [30].

LAPART adopts a neural inferencing mechanism to make predictions from learned patterns [30], [35]. LAPART architecture consist of two pattern classifier subnetworks (i.e., Fuzzy ART subnetworks referred to as the A-side and B-side) laterally connected with adaptive connections. The graphical illustration in **Figure 24** shows LAPART architecture with Fuzzy ART subnetworks [33]. This architecture implements a dual system of inference rules. Recognizing a member of a familiar category by one subnetwork prompts a rule which infers that a member of a familiar category will be recognized by the other subnetwork [30]. The second rule enables the other network to reject the prediction inconsistent with the input data. **Figure 25** illustrates the prediction rule in LAPART system. For example, if an input sample *x'* is found in '*category-a*', then the corresponding value *y'* should be in '*category-b*' if the inference is correct. The granularity (i.e., size of the hyperbox) of these categories in subnetwork-A and subnetwork-B are determined by vigilance parameters ρ_a and ρ_b respectively.

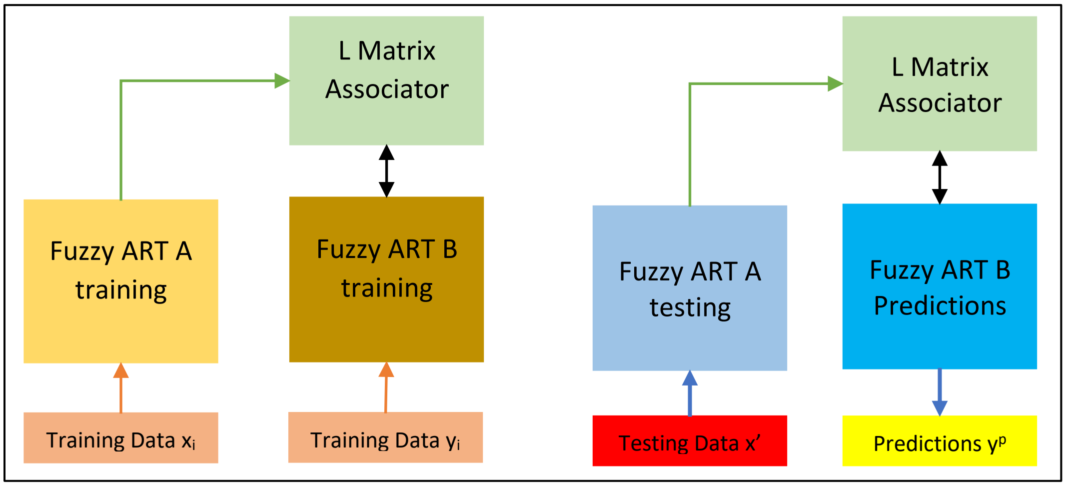


Figure 24: Graphical illustration of LAPART architecture

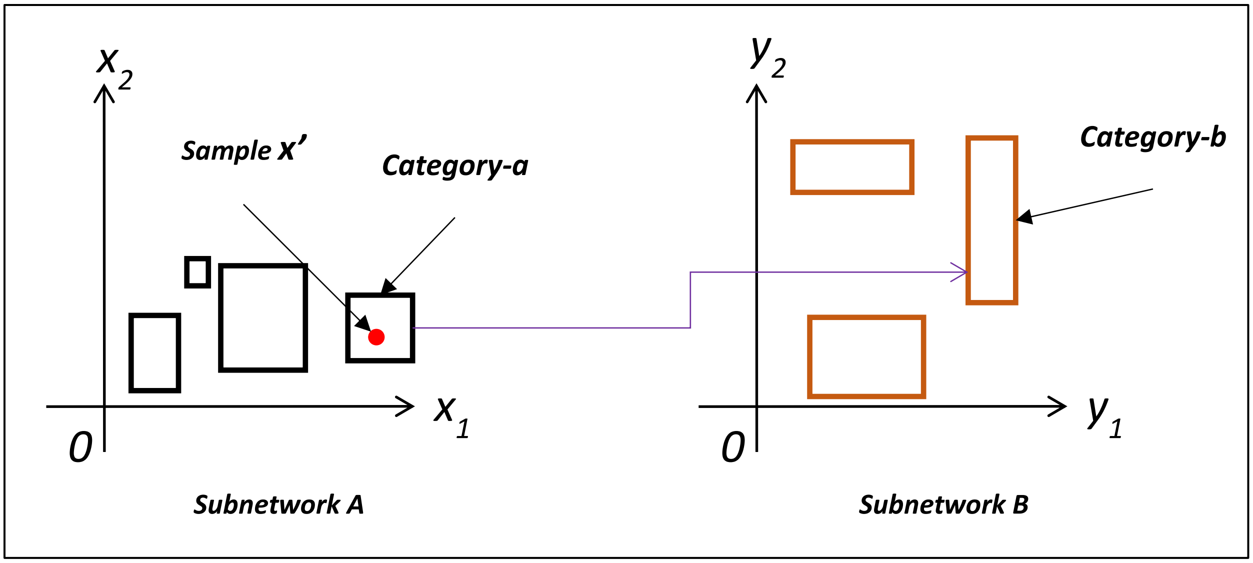


Figure 25: Associations between stimuli in LAPART

In this study, the training set for LAPART consists of N pairs of input patterns (i.e., N= 3x16 for lectures scheduled thrice a week and N=2x16 for lectures scheduled twice a week). The first member of the pair is for the A side input (from the 2D space) while the second member of the pair is for the B-side input (from the 9D space as shown in **Table 8**. The A-side input patterns consists of the first two measured occupancy samples of time-segments with dimensionality two and the B-side input patterns consists of the corresponding matching template profiles with dimensionality nine or twelve. The testing set has similar structure and dimensionality as the training set shown in **Table 8**. Here, LAPART system uses the A-side input patterns and the relations inferred during training to predict B-side template profiles. The B-side training set it then used to verify the predictions. The sample input data shown in **Table 8** corresponds to time-segment-1 scheduled from 10:00 am to 10:50am on Mondays, Wednesdays, and Fridays. Therefore, it has 48 patterns and **Table 8** shows the first 16 samples (i.e., 16 Mondays). Similar input datasets were formed for all the time-segments.

Table 8: Sample of LAPART Training Input

	A-Side Input		Template # for B-side inputs	B-side Input								
	A - 1	**A - 2**		**B - 1**	**B -2**	**B -3**	**B -4**	**B -5**	**B -6**	**B -7**	**B -8**	**B -9**
1	0.28	0.36	t1	0.21	0.31	0.43	0.45	0.47	0.47	0.47	0.46	0.21
2	0.19	0.33	t1	0.21	0.31	0.43	0.45	0.47	0.47	0.47	0.46	0.21
3	0.01	0.00	t2	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.06
4	0.19	0.30	t1	0.21	0.31	0.43	0.45	0.47	0.47	0.47	0.46	0.21
5	0.27	0.34	t1	0.21	0.31	0.43	0.45	0.47	0.47	0.47	0.46	0.21
6	0.06	0.26	t1	0.21	0.31	0.43	0.45	0.47	0.47	0.47	0.46	0.21
7	0.16	0.31	t1	0.21	0.31	0.43	0.45	0.47	0.47	0.47	0.46	0.21
8	0.12	0.29	t1	0.21	0.31	0.43	0.45	0.47	0.47	0.47	0.46	0.21
9	0.34	0.34	t1	0.21	0.31	0.43	0.45	0.47	0.47	0.47	0.46	0.21
10	0.09	0.41	t3	0.11	0.3	0.41	0.43	0.45	0.45	0.46	0.46	0.25
11	0.03	0.25	t3	0.11	0.3	0.41	0.43	0.45	0.45	0.46	0.46	0.25
12	0.25	0.25	t1	0.21	0.31	0.43	0.45	0.47	0.47	0.47	0.46	0.21
13	0.02	0.27	t3	0.11	0.3	0.41	0.43	0.45	0.45	0.46	0.46	0.25
14	0.03	0.25	t3	0.11	0.3	0.41	0.43	0.45	0.45	0.46	0.46	0.25
15	0.19	0.19	t3	0.11	0.3	0.41	0.43	0.45	0.45	0.46	0.46	0.25
16	0.06	0.22	t3	0.11	0.3	0.41	0.43	0.45	0.45	0.46	0.46	0.25

Since feasibility of occupancy profile prediction is being tested in this study, all the A-side samples from training were used in testing. LAPART system predicts a matching midpoint (average of template upper bound and lower bound) profile for every A-side input. These results provide evidence that occupancy profiles of a time-segment can statistically be distinguished by the first two measured occupancy samples of that time-segment.

3.3.Part-III: HVAC energy estimation

A comprehensive HVAC energy analysis was performed using the predicted occupancy (i.e., midpoint template profiles) for a 14-week period starting from August 18, 2019, to November 24, 2019. The energy simulations were performed in EnergyPlus, version 9.0. The simulations were performed on four different occupancy schedules: 1) fixed schedule, 2) registered schedule, 3) WiFi schedule, and 4) predicted schedule. The fixed/static schedule assumes maximum occupancy for the lecture hall and the HVAC system runs from the beginning of the first time-segment in the day to the end of last time-segment in the day. The registered schedule runs the HVAC systems only during the time-segments (i.e., scheduled lectures) assuming registered students per lecture as its occupancy. The WiFi schedule runs the HVAC systems using the actual measured occupancy during the time-segments. Additionally, the occupancy measured outside the durations of time-segments were omitted from the WiFi schedule. The predicted schedule runs the HVAC systems using predicted composite profiles (i.e., concatenating the midpoint profiles predicted for individual time-segments in Section 3.2).

The lecture hall is served by a dual-duct HVAC unit and hence, a dual-duct model was built in EnergyPlus for all the energy consumption estimates in this study. The dual-duct model was designed using the instructions provided in EnergyPlus 'Input Output Reference'. The modules that are essential for a dual-duct model as listed in EnergyPlus – Input Output Reference manual are: 1) HVAC Template: Thermostat, 2) HVAC Template Zone: Dual Duct, 3) HVAC Template: Plant – Chilled water loop, 4) HVAC Template: Plant - Chiller, 5) HVAC Template Plant: Tower, 6) HVAC Template: Plant Hot water loop, and 7) HVAC Template: Plant – Boiler. **Figure 26** shows the list of modules created for the dual-duct HVAC system in EnergyPlus model.

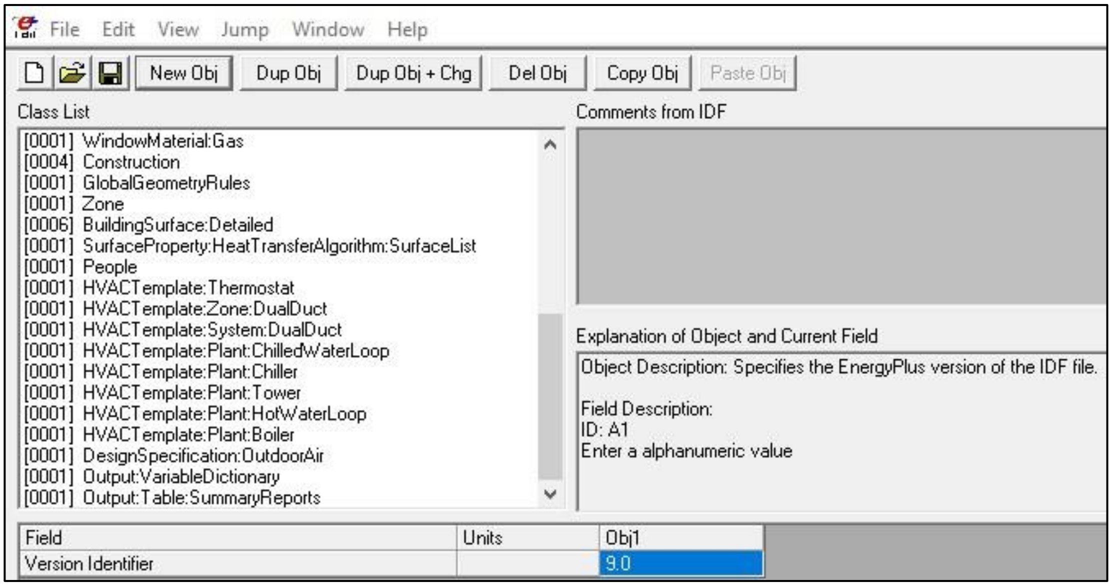


Figure 26: EnergyPlus Dual-Duct model's module list

In addition to the list of modules shown in **Figure 26**, the occupancy schedules were added to 'Schedule Compact' module. Additionally, thermostat schedules for setpoint temperature were created for corresponding occupancy schedules. These EnergyPlus schedules were built for a 14-week period starting from August 18, 2019, to November 24, 2019. The limitations within EnergyPlus Schedule Compact module restricted the entries for occupancy schedule to a 14-week period. However, a 14-week period should be adequate enough to draw conclusions on energy consumption for different occupancy schedules.

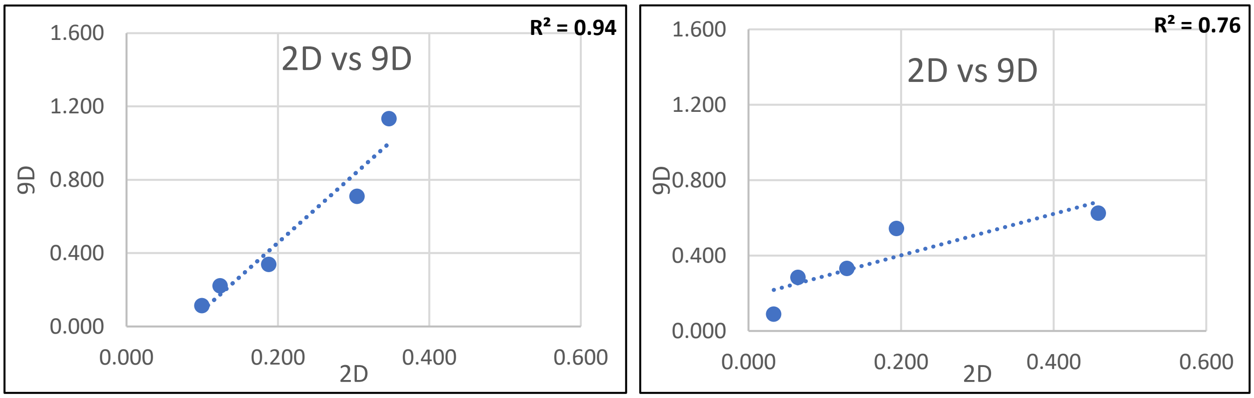
4. Results

The HVAC energy consumption values for four different occupancy conditions listed in Section 3.1 are provided in **Table 9**.

Table 9: HVAC energy consumption for a 5-day week

Occupancy Type	Energy Consumed [GJ]
Fixed Schedule	2.44
Registered Schedule	1.28
WiFi Schedule	1.04
MLT- Schedule	0.84

The pairs of multi-dimensional points (i.e., 2-D, and 9-D or 12-D) calculated per template w.r.t the MLT of each time-segment were plotted as shown in **Figure 27**. The example correlation plots provided in **Figure 27** are for time-segments #1, #2, #3, and #4 where time-segments #1, #2, and #3 are 2-D vs 9-D plots and time-segment #4 is a 2-D vs 12-D plot. The r^2 values for each time-segment ranged between 0.76 to 0.98.



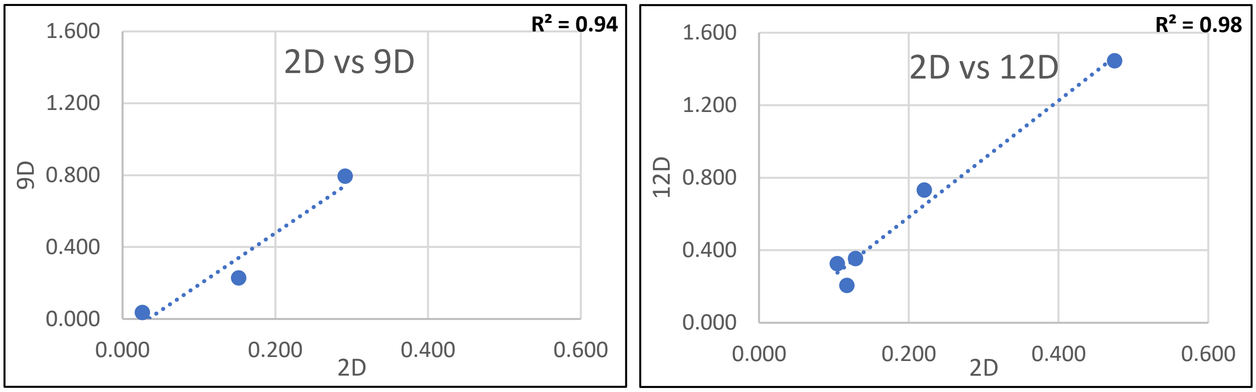


Figure 27: Correlation between A-side and B-side inputs for different time-segments. Top row: time-segment #1 & #2, bottom row: time-segment #3 & #4.

Figure 28 shows LAPART predicted midpoint occupancy profile (shown in green) at $\rho_a = 0.90$ and $\rho_b = 0.90$. **Figure 28** includes the A-side input of two measured occupancy samples (shown in black), and the corresponding B-side input of average template profile (shown in red) for time-segment #1. The R^2 value for predicted midpoint profile and input average template profile in **Figure 28** is 0.955. **Figure 29** shows four predicted midpoint occupancy profiles that matched with the 48 A-side inputs for time-segment #1 at $\rho_a = 0.90$ and $\rho_b = 0.90$. Similarly, midpoint profiles for the given A-side inputs were predicted for all the ten time-segments at $\rho_a = 0.90$ and $\rho_b = 0.90$. **Table 10** provides the training error for LAPART algorithm estimated for all time-segments.

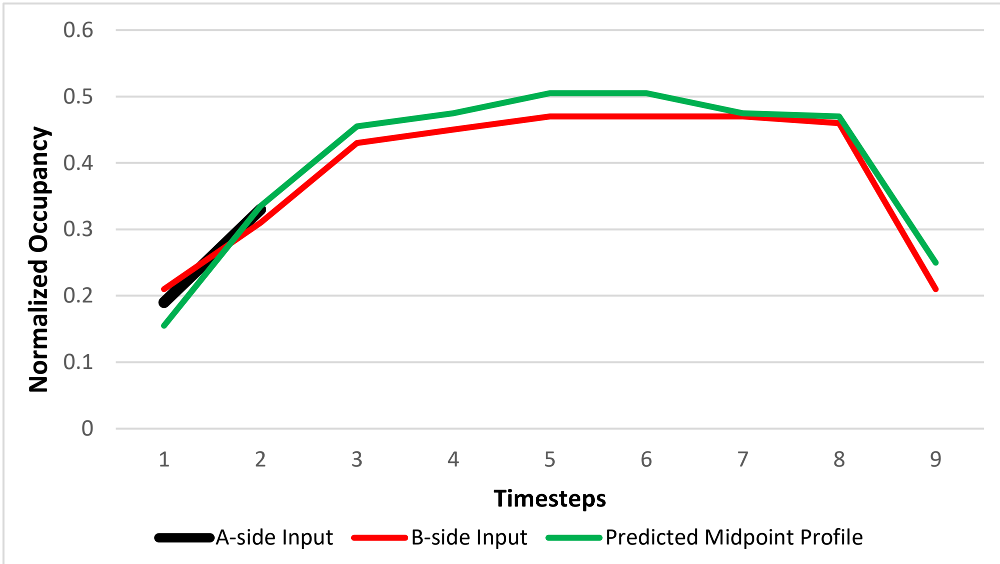


Figure 28: LAPART predicted occupancy profile for a given A-side input.

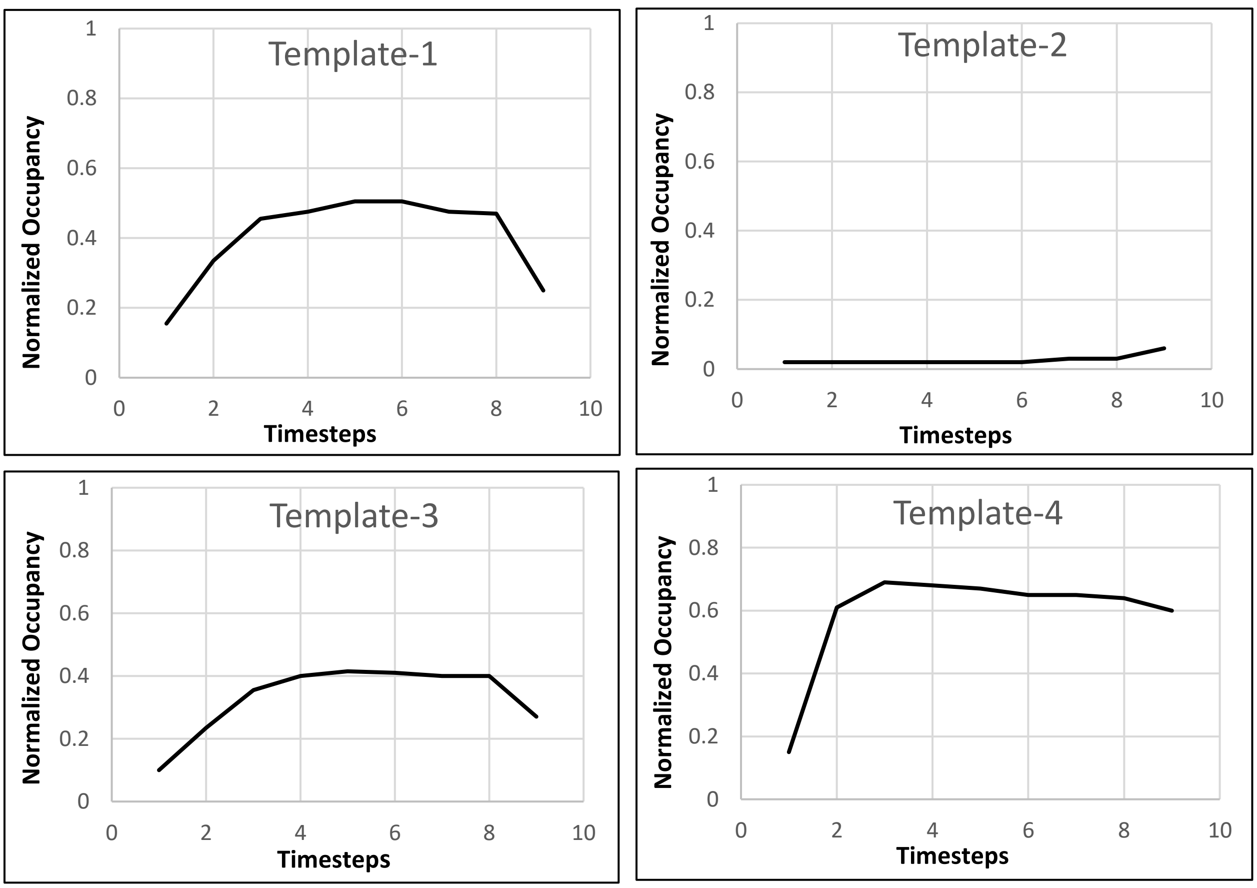


Figure 29: LAPART predicted midpoint template profiles for time-segment #1 at A-ρ = 0.90 and B-ρ = 0.90

Table 10: Training Error for LAPART

Training Error	
Time-Segment	**Error Fraction**
Time-Segment -1	0.063
Time-Segment -2	0.021
Time-Segment -3	0.000
Time-Segment -4	0.000
Time-Segment -5	0.000
Time-Segment -6	0.000
Time-Segment -7	0.000
Time-Segment -8	0.000
Time-Segment -9	0.000
Time-Segment -10	0.000

Figure 30 shows an example 24hr. composite profile formed for Monday (week-1) using the predicted midpoint profiles for time-segments 1 to 4. Similar composite profiles were created for each weekday for all 16-weeks. These composite profiles were used as occupancy for the 'predicted schedule' in EnergyPlus HVAC model as explained in Section 3.3. EnergyPlus simulations were performed for four different occupancy schedules listed in Section 3.3. The energy consumption was estimated for a 14-week period from August 18, 2019, to November 24, 2019. The 14-week energy consumption for all the four occupancy schedules listed in Section 3.3 are presented in **Table 11**.

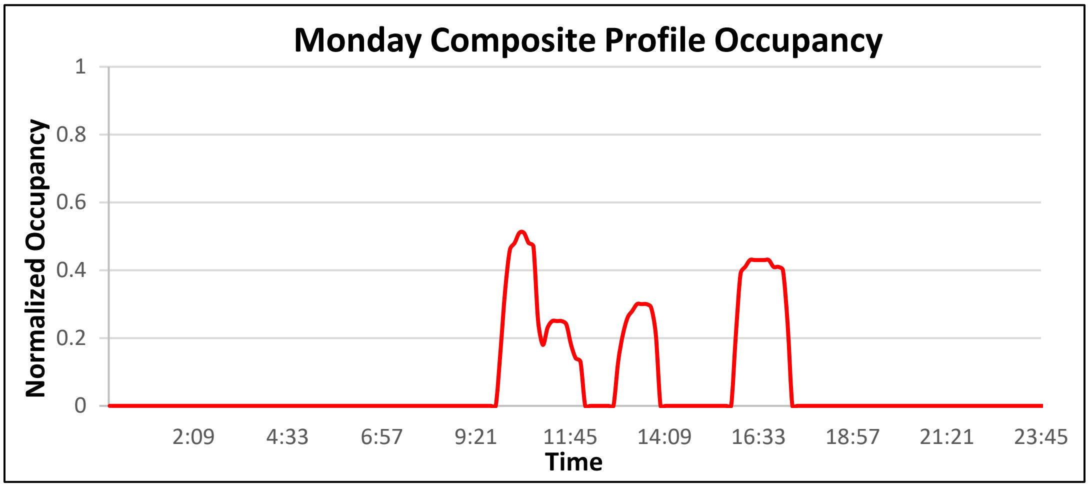


Figure 30: A 24hr composite profile created from predicted midpoint profiles for time-segments 1 to 4.

Table 11: HVAC energy consumption for a 14-week period

Occupancy Schedule Type	Total Energy [GJ]
Fixed occupancy	43.92
Registered occupancy	18.30
WiFi occupancy	14.51
Predicted occupancy	14.27

5. Discussion

This section discusses the implications, potential benefits, and limitations of this study. The proposed methodology has the following implications.

The current occupancy prediction strategies use complex occupancy data and unique IDs to make predictions. Though these methods are successful, their applicability is limited to buildings that have occupants who can be categorized as ‘long-term’ occupants. In this study, the patterns learned from WiFi measured occupancy data provide an alternative way to forecast occupancy in buildings with occupants that cannot be classified as ‘long-term’ occupants. From the patterns learned, the MLT pattern identified for each time-segment can serve as an occupancy forecast for that time-segment. The HVAC energy consumption results for different occupancy schedules presented in **Table 9** justify the use of MLT patterns as a baseline for expected occupancy. As explained in Section 3.2, MLTs may not always match with the actual occupancy of a time-segment.

To address this, a variable that has sufficient information to predict a better matching template is needed. The variable examined in this study was the pair of time-samples at the beginning of each time-segment. The correlation analysis and the results presented in **Figure 27** suggests that the first two measured occupancy samples of a time-segment can possibly be used to predict a better matching template for that time-segment. Based on this premise, occupancy template predictions were made for a time-segment using the first two measured occupancy samples. The predicted profiles are shown in **Figure 29** and it can be observed that the first two measured samples contain sufficient information to predict templates with statistical variability. For example, Template -2 and Template-4 illustrate statistical variability compared to Template-1 and Template-3. A visual comparison between the predicted profile and the input profiles shown in **Figure 28** highlight the prediction accuracy. Additionally, the low training error presented in **Table 10** further strengthens the prediction accuracy.

The modeled HVAC energy consumption results are provided in **Table 11**. As expected, the 14-week energy consumption for the fixed occupancy schedule is the highest. The energy consumption for registered occupancy suggests that when other source of occupancy data is available, a registered schedule can be used to realize significant energy reduction compared to a fixed occupancy schedule. Furthermore, the energy consumption of WiFi occupancy highlights that the actual occupancy is typically lower than registered occupancy emphasizing the importance of collecting occupancy data. The energy consumption of Predicted Occupancy is within 2% of WiFi occupancy's energy consumption that further validates the prediction occupancy.

The proposed methodology for occupancy forecast uses existing infrastructure that eliminates additional investments and occupant privacy concerns. More importantly, this approach can be implemented in commercial buildings such as shopping centers, airports, and gymnasiums where the occupants cannot be designated as 'long-term' occupants. The MLT patterns identified from the set of patterns learned for a time-segment can be used as a baseline for occupancy forecast of that time-segment. Additionally, the results lend evidence to the possibility of using variables that are significantly correlated to the occupancy, a better matching template can be predicted for a time-segment.

Knowing the occupancy count can be used to precondition a space/zone ahead in time to increase occupant thermal comfort during a time-segment. As the impact of occupants' metabolic rates and the associated heat loss is evident in the literature, the temperature setpoints can be made a function of expected number of occupants. This temperature setpoint schema accounts for the total heat loss from n number of occupants and estimates a new setpoint T_n that is lower/higher (i.e., lower for summer months and higher for winter months) than the standard temperature setpoints. For example, different temperature setpoints can be employed for different occupant counts. This particular practice not only ensures the occupant thermal comfort but also reduces the HVAC demand (in case of low occupancy). Additionally, as the metabolic rates vary throughout the day, different setpoint can be used for different time-segments with similar occupancy. For examples, T_{n1} can be the setpoint for time-segment#1 with 40 occupants that takes place at 9am and T_{n4} can be the setpoint for time-segment#4 with 40 occupants that takes place at 1:00pm. This temperature setpoint schema can positively impact occupants' thermal comfort and their health.

The preconditioning can be advantageous in a scenario where the time-lag to bring the space temperature to the setpoint is high. If the time-lag is smaller, real-time occupant information is sufficient to regulate the temperature of the space.

6. Summary and Conclusion

The feasibility test conducted in this study highlights that the intrinsic variables with significant correlation with occupancy can be used as a predictor for occupancy pattern (via learned templates) of a time-segment. This conclusion can be extended to extrinsic variables that significantly impact occupancy of a time-segment.

In summary, significant patterns can be learned from WiFi data that does not contain any form of occupant identifiers. Occupancy predictions are possible from the learned patterns for a relative dynamic environment such as airports, libraries, institutional buildings, and shopping malls. As occupants in such commercial spaces cannot be classified as 'long-term' residents, the current framework provides a novel occupancy prediction strategy. Additionally, the current approach of using the first two measured samples as predictors for the matching occupancy profile can aid in minimizing wasteful HVAC demand and avoiding occupant thermal discomfort. In future, the current methodology and lessons learned can be implemented for different types of commercial buildings that do not have a registered occupancy and external schedule (e.g., shopping malls, airports, and gymnasiums). External variables such as deals offered at a warehouse (e.g., Costco), flights scheduled at an airport, weather related flight schedule changes/cancellations, and gymnasium memberships can have significant correlation with the occupancy. These variables can be examined to find their potential in forecasting occupancy count. Furthermore, a detailed occupancy comfort study can be conducted to quantify the impact of the proposed variable temperature setpoint schema.

Acknowledgment

The study acknowledges the contributions made by the IT department and the Physical Plant department of University of New Mexico.

References

[1] ASCE, "Infrastructure Report Card," no. March, pp. 1–74, 2013.

[2] EIA and B. Hojjati, "Global energy consumption driven by more electricity in residential, commercial buildings," *U.S. Energy Inf. Assoc.*, vol. 2019, pp. 2019–2020, 2019.

[3] C. Basu, G. Ghatikar, and P. Bansal, "Enabling Efficient, Responsive, and Resilient Buildings: A Collaboration Between the United States and India," no. March, 2014.

[4] Y. Jiang, D. Yan, S. Guo, and S. Hu, "China Building Energy Use 2018," no. December 2016, 2019.

[5] W. Wang, J. Chen, T. Hong, and N. Zhu, "Occupancy prediction through Markov based feedback recurrent neural network (M-FRNN) algorithm with WiFi probe technology," *Build. Environ.*, vol. 138, no. March, pp. 160–170, 2018.

[6] E. Azar and C. C. Menassa, "A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings," *Energy Build.*, vol. 55, pp. 841–853, 2012.

[7] J. Dong, C. Winstead, J. Nutaro, and T. Kuruganti, “Occupancy-based HVAC control with short-term occupancy prediction algorithms for energy-efficient buildings,” *Energies*, vol. 11, no. 9, pp. 1–20, 2018.

[8] Z. Chen, C. Jiang, and L. Xie, “Building occupancy estimation and detection: A review,” *Energy Build.*, vol. 169, pp. 260–270, 2018.

[9] Y. Peng, A. Rysanek, Z. Nagy, and A. Schlüter, “Using machine learning techniques for occupancy-prediction-based cooling control in office buildings,” *Appl. Energy*, vol. 211, no. November 2017, pp. 1343–1358, 2018.

[10] K. C. J. Simma, S. M. Bogus, and A. Mammoli, “WI-FI ROUTER NETWORK-BASED OCCUPANCY ESTIMATION TO OPTIMIZE HVAC ENERGY CONSUMPTION,” vol. 2015, no. Saic 2011, pp. 1–10, 2019.

[11] K. C. J. Simma, A. Mammoli, and S. M. Bogus, “Real time occupancy estimation using WiFi network to optimize HVAC operation,” vol. 00, no. 2018, 2019.

[12] W. Wang, J. Chen, and T. Hong, “Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings,” *Autom. Constr.*, vol. 94, no. July, pp. 233–243, 2018.

[13] Y. Ding, W. Chen, S. Wei, and F. Yang, “An occupancy prediction model for campus buildings based on the diversity of occupancy patterns,” *Sustain. Cities Soc.*, vol. 64, no. October 2020, 2021.

[14] J. Wang, N. C. F. Tse, and J. Y. C. Chan, “Wi-Fi based occupancy detection in a complex indoor space under discontinuous wireless communication: A robust filtering based on

event-triggered updating," *Build. Environ.*, vol. 151, no. November 2018, pp. 228–239, 2019.

[15] W. Wang, J. Chen, G. Huang, and Y. Lu, "Energy efficient HVAC control for an IPS-enabled large space in commercial buildings through dynamic spatial occupancy distribution," *Appl. Energy*, vol. 207, pp. 305–323, 2017.

[16] W. Parys, D. Saelens, and H. Hens, "Coupling of dynamic building simulation with stochastic modelling of occupant behaviour in offices - a review-based integrated methodology," *J. Build. Perform. Simul.*, vol. 4, no. 4, pp. 339–358, 2011.

[17] H. B. Gunay, W. O'Brien, and I. Beausoleil-Morrison, "A critical review of observation studies, modeling, and simulation of adaptive occupant behaviors in offices," *Build. Environ.*, vol. 70, pp. 31–47, 2013.

[18] S. S. K. Kwok, "A study of the importance of occupancy to building cooling load in prediction by intelligent approach," *Energy Convers. Manag.*, vol. 52, no. 7, pp. 2555–2564, 2011.

[19] S. H. Ryu and H. J. Moon, "Development of an occupancy prediction model using indoor environmental data based on machine learning techniques," *Build. Environ.*, vol. 107, pp. 1–9, 2016.

[20] N. Li, G. Calis, and B. Becerik-Gerber, "Measuring and monitoring occupancy with an RFID based system for demand-driven HVAC operations," *Autom. Constr.*, 2012.

[21] M. S. Zuraimi, A. Pantazaras, K. A. Chaturvedi, J. J. Yang, K. W. Tham, and S. E. Lee, "Predicting occupancy counts using physical and statistical Co2-based modeling methodologies," *Build. Environ.*, 2017.

[22] D. Li, C. C. Menassa, and V. R. Kamat, "Personalized human comfort in indoor building environments under diverse conditioning modes," *Build. Environ.*, vol. 126, no. September, pp. 304–317, 2017.

[23] Q. Wang, G. Augenbroe, J. H. Kim, and L. Gu, "Meta-modeling of occupancy variables and analysis of their impact on energy outcomes of office buildings," *Appl. Energy*, vol. 174, pp. 166–180, 2016.

[24] K. Ahmed, J. Kurnitski, and B. Olesen, "Data for occupancy internal heat gain calculation in main building categories," *Data Br.*, vol. 15, pp. 1030–1034, 2017.

[25] W. D. Van Marken Lichtenbelt and B. R. Kingma, "Building and occupant energetics: A physiological hypothesis," *Archit. Sci. Rev.*, vol. 56, no. 1, pp. 48–53, 2013.

[26] J. Mege, "Variations , and the Immune System," 2021.

[27] L. Klein *et al.*, "Coordinating occupant behavior for building energy and comfort management using multi-agent systems," *Autom. Constr.*, vol. 22, pp. 525–536, 2012.

[28] C. Liao and P. Barooah, "An Integrated Approach to Occupancy Modeling and Estimation in Commercial Buildings," pp. 3130–3135, 2010.

[29] G. A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system," *Neural Networks*, vol. 4, no. 6, pp. 759–771, 1991.

[30] M. J. Healy, T. P. Caudell, and S. D. G. Smith, “A neural architecture for pattern sequence verification through inferencing,” *IEEE Trans. Neural Networks*, vol. 4, no. 1, pp. 9–20, 1993.

[31] M. J. Healy and T. P. Caudell, “Acquiring rule sets as a product of learning in a logical neural architecture,” *IEEE Trans. Neural Networks*, vol. 8, no. 3, pp. 461–474, 1997.

[32] G. A. Carpenter, S. Grossberg, and D. B. Rosen, “Fuzzy ART: An adaptive resonance algorithm for rapid, stable classification of analog patterns,” *Proceedings. IJCNN - Int. Jt. Conf. Neural Networks*, no. August 1991, pp. 411–416, 1992.

[33] C. B. Jones, “Fault detection and diagnostics of an HVAC sub-system using adaptive resonance theory neural networks,” 2015.

[34] K. C. J. Simma, T. P. Caudell, S. M. Bogus, and A. Mammoli, “FUZZY ART : PATTERN RECOGNITION OF WIFI DETECTED OCCUPANCY IN COMMERCIAL BUILDINGS,” *Springer, Nat.*, no. May, 2021.

[35] T. P. Caudell and M. J. Healy, “Lateral priming adaptive resonance theory (LAPART)-2: innovation in ART,” in *Recent Advances in Artificial Neural Networks*, CRC Press, 2018, pp. 187–211.

Chapter-6: Conclusion

6.1. Overall Conclusion

The results from the experiments and analyses conducted in this research contribute to a better understanding of occupancy structure, patterns, and the possibility of occupancy prediction in commercial buildings. Chapter 2 of this dissertation established that reliable occupancy data can be obtained from WiFi networks that are free from all occupant identifiers. The device count obtained from pre-existing WiFi routers consistently measured the actual occupancy (i.e., ground truth obtained using people counting sensors) of a lecture hall for a 16-week period. Typically, the device count data is passively collected by WiFi routers and does not require upgrades or added infrastructure. The occupant identifiers such as MAC addresses or IP addresses are not needed to estimate the occupancy thereby reducing occupant's privacy concerns. This simplistic way of obtaining occupancy that does not require additional investment or raise occupancy privacy concerns can be favored by building owners and their efforts towards building energy reduction.

Chapter 3 investigated the effects of different occupancy schedules on HVAC energy consumption. The results in Chapter 3 provide evidence that real-time occupancy data are required to maximize HVAC energy savings. On the contrary to the existing literature, occupants and their behavior remained unaltered during data collection period allowing to capture true occupant dynamics from a relatively complex environment. In addition to the occupancy schedules based on WiFi data, occupancy schedules were built using registered occupancy data (i.e., data of registered students per class) and people counter data (i.e., ground truth). In general, occupancy-based schedules resulted in significant HVAC energy reduction compared to a fixed/static schedule that typically assumed maximum occupancy for a space. When available, registered occupancy can result in significant HVAC energy savings, but the real-time occupancy data can further increase the energy savings. These results emphasize the need for unaltered real-time occupancy data to reduce HVAC energy consumption in commercial buildings.

Chapter 4 investigated the structure of WiFi measured occupancy data to identify patterns of repetitions using neural networks. Three different experiments were conducted with the data collected between August 2019 and December 2019. The experiments were conducted using different combinations of datasets and Fuzzy ART neural network. The results from these experiments suggest that significant pattern recognition is possible by dividing occupancy profiles into smaller segments bound by external schedules. These smaller segments are termed as time-segments and the external schedules are typically the lecture durations, rush hours, and peak hours. Additionally, from the set of occupancy patterns learned for each time-segment, at least one pattern repeats frequently which is termed as Most Likely Template (MLT). Furthermore, each time-segment may also have a Holiday Template (HT) that represents zero-occupancy profiles for the time-segment. The MLT occupancy profile learned for each time-segment can be treated as expected occupancy for those time-segments based on which the HVAC systems can be operated.

Chapter 5 investigated the feasibility of occupancy prediction using the occupancy data that is free of occupant identifiers such as MAC address, IP addresses, and unique IDs. Using the occupancy patterns learned and the LAPART neural network, occupancy prediction analysis was conducted. The results from the analysis suggest that occupancy profile prediction for a time-segment is possible if there exists a variable (i.e., a predictor) that is significantly correlated to the occupancy profile of that time-segment. Contrary to the existing literature that uses unique occupant identifiers to predict their presence in a space and consequently the space occupancy, this approach predicts a matching occupancy profile for a time-segment using a predictor. The feasibility tests conducted in this study used the first two measured occupancy values (i.e., intrinsic variables) to predict a matching profile from the set of profiles learned for each time-segment. As suggested previously, the MLT profiles can be used as expected occupancy. However, when the actual occupancy deviates from the MLT profile, using the first two measured occupancy values to predict a matching profile can avoid incorrect occupancy profile prediction. This mechanism ensures that there is no unwanted HVAC demand and insufficient thermal conditioning when occupancy is overestimated and underestimated, respectively. Similarly, occupancy profile prediction can be made using extrinsic variables if they are significantly correlated to the occupancy profiles.

Using the occupancy count, a new strategy for indoor thermal setting can be implemented. The indoor heating and cooling requirements are dependent on occupant homotherm nature and their thermal interactions with indoor climate. Therefore, dynamic heating and cooling setpoints that are a function of occupant count (e.g., a lower cooling setpoint for higher occupancy relative to standard practice that accounts for occupant heat radiation at a given time of the day) can aid in increasing occupant thermal comfort inside buildings. A forecast of occupant count can allow for preconditioning a space using the proposed dynamic thermal setpoints. As the preconditioning temperatures are dependent on the expected occupant count, the time required to pre-condition the space changes with the occupant count. This emphasizes the need for an occupant count forecast to ensure maximum occupant comfort and an unwanted HVAC demand.

While the current framework eliminates the occupant privacy concerns, the use of WiFi networks for occupant data highlights the need for strengthening the cybersecurity against attacks. When the proposed framework is implemented to regulate indoor climate control of commercial buildings, the interdependency between the WiFi network and the HVAC systems increases. The interdependency between the WiFi and HVAC infrastructure becomes critical and a cyber-attack on the WiFi network can adversely affect the occupant comfort inside buildings. Therefore, strong security measures are needed for protecting the WiFi networks prior to its interaction with the HVAC infrastructure. Standards such as ANSI/TIA-862, ANSI/ASHRAE Standard 135-1995, BACnet, RESTful web services, and WPA2 among other can be implemented for securing the wireless connections and IoT devices on the networks [1].

6.2. Implications and Future Directions

This research implies that WiFi networks in commercial buildings provide a reliable source of occupancy data that can be used to create demand driven HVAC operational schedules to reduce building energy consumption. Contrary to the existing literature, the identifier free occupancy data collected from a relatively dynamic environment represents occupancy of large commercial spaces such as airports, libraries, universities, and other public buildings. Therefore, the energy consumption results in this research signify the energy saving potential in large commercial buildings. Results from pattern analysis imply that patterns of occupied and unoccupied segments can be detected from dynamic environments using WiFi data. Furthermore, the occupancy profile prediction results lend evidence to the possibility of occupancy load forecasting that can aid effective HVAC operation to reduce energy consumption and increase occupant comfort.

In the future, a study of extrinsic variables that are significantly correlated to the occupancy could be conducted. Furthermore, the feasibility of occupancy prediction ahead of time using extrinsic variables could be investigated. The occupancy forecast could be used to preheat and precool the space ahead of scheduled occupancy using the proposed dynamic thermal setpoints. The effect of dynamic thermal setpoints for preheating and precooling a space on occupant comfort levels could then be investigated. This research took advantage of an available occupancy schedule for the lecture hall to identify occupied and unoccupied segments. Occupancy data from different types of commercial buildings that do not have an occupancy schedule could be analyzed to identify novel ways to detect the occupied and unoccupied segments. Additionally, the patterns from different building types could be compared and their temporal variability could be investigated to understand the pattern stability over time.

References:

[1] R. Khatoun and S. Zeadally, "Cybersecurity and privacy solutions in smart cities," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 51–59, 2017.