Occupancy matters:

Toward an occupancy-driven ventilation system

Using WiFi infrastructure and Neural Network

Nazanin Abbaszadeh Bajgiran

A thesis

presented to Ryerson University

in fulfillment of the

thesis requirement for the degree of Master of Science

In the program of

Computer Networks

Toronto, Ontario, Canada, 2017

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

Abstract

Buildings accounted for 20% of total energy consumption and 54% of electricity usage in 2013 in Canada. Heating Ventilation and Air conditioning system is the main consumer of energy in the buildings. The common approach for designing a ventilation system is a predefined schedule based on the maximum capacity disregard the actual number of occupants. We believe that passive use of already existing WiFi infrastructures can replace the monitoring sensors and cut the cost of energy and extra sensors installation. A field study was conducted in graduate offices of Ryerson University to examine the opportunity of energy saving by changing the fixed ventilation schedule to the occupancy driven one. The number of occupants had been determined using pre-existing WiFi infrastructure and by using the real time occupancy data, the new system achieved 76% reduction in ventilation energy consumption.

To further investigate the potentials of WiFi infrastructure, Finger Printing and Neural Network method had been used to map the occupant's location by analyzing the Received Signal Strength Indicator (RSSI) of the wifi equipped device. The results showed 95% accuracy in the first round of testing and 92% accuracy after 1 week of retesting the model by using pattern recognition technique. Employing this approach could lead to even more energy saving by assigning the required airflow to each subzone proportionally to the number of its occupants.

Acknowledgements

I would like to express my gratitude to my professor Dr. Kaamran Raahemifar for his support, guidance and patience throughout my research work. His knowledge and experience added extra value to my thesis.

Special thanks to Dr. Bobby Ma for his continuous support since I joined Ryerson University.

Words cannot express the amount of love and gratitude I have for my beloved family. My lovely mother and father, I cannot thank you enough for the sacrifices you made. If it was not for your support, I could not be where I am right now. I am forever thankful to my lovely sister and brothers for their emotional support and encouragements from thousand miles away.

I would like to express my gratefulness to Mr. Joshua Cave form CCS department of Ryerson, for his expertise and time. He patiently provided me with the requested data and was always helpful. I wish to appreciate my friend Ali for sharing his technical knowledge and experience on my thesis. I would like to thank my friend Taha for sharing his knowledge and motivation. I also wish to thank my friend Marjan who helped me a lot emotionally and was always pushing me forward.

At the end, I would like to thank Mehdi, for his support, without his contribution this thesis would not have become fruitful.

Dedication

This thesis is dedicated to my father for his sacrifices and trust in my abilities and to my mother for her unconditional love and support.

My Sister and my brothers for always being there for me and helped me to survive familial ties half across the world.

Table of Contents

List of Figures

LIST OF TABLES

Chapter 1. Introduction

1.1 General information

According to National Resources Canada, buildings (residential, commercial) are responsible for the consumption of 20% of total energy and about half of total electricity in 2013 [1].

Retrofitting existing buildings, energy conservation policies and making use of renewable energy has gained more interest in recent years. This is due to the rise in global energy consumption, energy cost, greenhouse gas (GHG) emission and climate change.

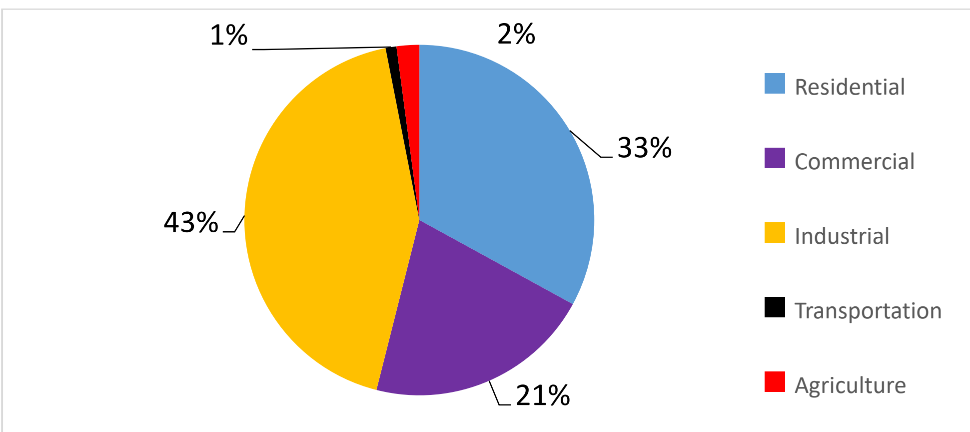

Figure 1- Electricity energy use by sector, Canada, 2013 [1].

The factors that affect the energy consumption of a building can be categorized into six parts:

1- Climate
2- Building features (e.g., location, type, etc.)
3- Building systems (e.g., HVAC)

4- Occupants' behavior and activities

5- Social and economic factors (e.g., education, income)

6- Indoor quality

Social and economic factors mostly influence the occupants' behavior towards the energy usage. [2]

Between 1990 and 2013 energy efficiency has been improved by 24% and contributed to a pronounced reduction in the energy demand rate. In fact, energy load increased by 28%, but this number would have increased by 51% without energy efficiency improvements (Figure 2) [1].

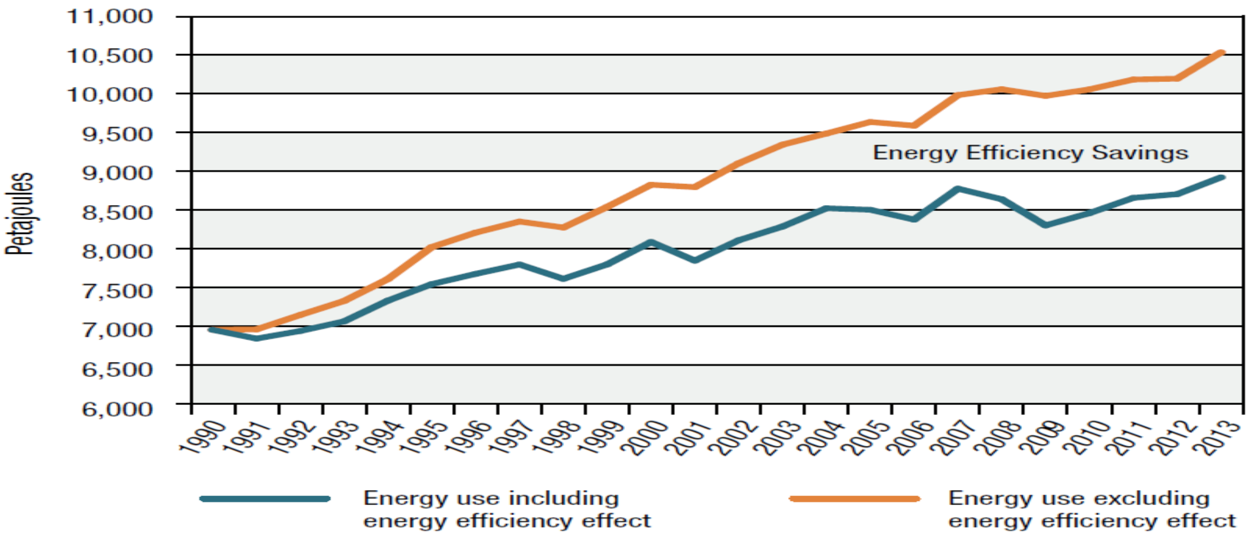

Figure 2- energy use in Canada with and without efficiency improvements [1]

Contrary to popular belief, energy conservation methods are not expensive and they do not interfere with the comfort of the consumers. In fact, the retrofitting investment would be returned after a while along with business advancement.

Energy demand is different in residential and commercial buildings. Space heating, lighting and equipment are the major factors in commercial sectors, while space and water heating are the main end users of energy in residential buildings (Figure 3).

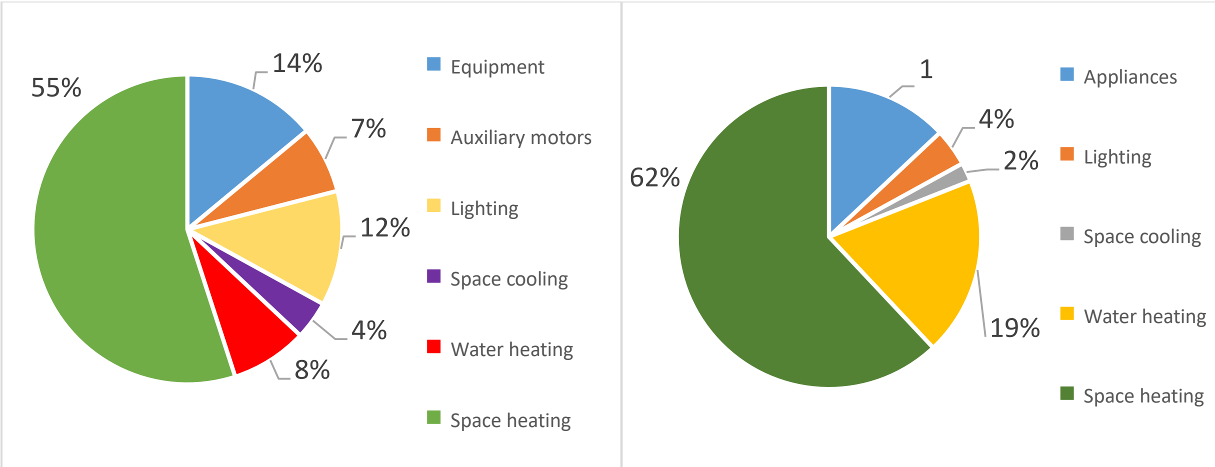

Figure 3- Commercial electricity demand (left), Residential electricity demand (right) [1]

Heating, Ventilation and Air Conditioning system (HVAC), lighting and IT equipment are serving occupants to provide their comfort. That is why building energy consumption is heavily affected by its occupants. Therefore, knowledge of building occupancy is crucial to reduce energy consumption [3]. Occupancy pattern can be used by building management system (BMS) to react proportionally to the number of occupants in order to save energy. It has been investigated that turning off the HVAC system when there is no occupant inside could result in 30% reduction of this energy [4].

Different approaches have been proposed and implemented to detect occupancy in different types of buildings. Sensors such as RFID tags, CO_2 meters, motion detectors, cameras, etc. have been implemented in previous studies. Nevertheless, these approaches required extra investments and also may violate occupants' privacy. Moreover, most of these sensors make a binary decision of occupied or unoccupied rather than the real number of occupants.

Commonly ventilation systems are operating on a predefined schedule basis and for the maximum occupancy, while there are rooms in the building that may have a few or none occupant at certain times. By detecting the number of occupants and their zone, this thesis aims to reveal the extent of occupancy impact on energy saving.

1.2 Thesis objectives

Most of the previous occupancy detection studies were not seeking energy saving and only small portion of those with the goal of energy saving (e.g. the occupancy based HVAC systems and lighting systems) have been tested in real life scenarios.

Most of these studies suggest changing the airflow based on a prediction of peak hours and non-peak hours and not the real time of occupancy. The occupancy driven HVAC systems were often leveraging extra hardware that came with extra costs and maintenance.

Consequently, the objectives of this thesis are to investigate the following:

1) The feasibility of using existed WIFI access points to detect the presence of occupants in a building.
2) The feasibility of using existed access points to count the number of local occupants in a building housing single occupant or multi occupants offices. The local occupant is denoted as the one who stays in one place for a while i.e. more than five minutes and not just passing by.
3) Managing the ventilation system based on the real number of occupants.
4) Energy consumption reduction based on a total number of occupants.
5) Zone detection of the occupant using WIFI access points and neural network to adjust the ventilation system based on the demand of sub-zones.

   The idea is to collect data for three different days. Two days for training data set and the third one for testing as we wanted to check the stability and accuracy degradation of our system.

1.3 Thesis Organization

This thesis has been organized as follows: Chapter 1 presents an introduction to the thesis topic and the objectives of the research. Chapter 2 reviews the literature on occupancy monitoring systems, ventilation systems, indoor localization and neural networks. In chapter 3, experiment setup and methodology is discussed in two major parts. The first part uses access point data to calculate the population of the target zone to adjust ventilation based on occupancy. The second part uses received signal strength to detect the indoor location of the user in order to examine the opportunity of even more energy saving by allocating variable ventilation share to different sub-zones. Conclusion and future works are presented in chapter 4.

Chapter 2. Literature review

HVAC is the dominant energy consumer in the buildings. It accounts for 43% of energy consumption in the US and 61% in Canada which has a colder climate [1]. Previous literature shows that a huge number of modern buildings use the predefined static schedule to control HVAC system [5,6,7]. Recently, new buildings utilize Variable Air Volume (VAV) to control each zone separately. However, in practice due to lack of the real data of the occupancy, it is not being employed efficiently. Although Heating and Ventilation consume almost the same amount of energy, the number of studies regarding ventilation energy consumption is much less than the heating part. Ventilation system is responsible to maintain a good level of indoor air quality (IAQ) by circulating the fresh air in different zones of the building. As occupants spend more than 90% of their time indoors the IAQ will affect their comfort, health and productivity [8]. Thus, none dynamic ventilation system not only results in energy waste, but it could endanger occupant's health as well. Different sensors and methods had been used to investigate the feasibility of occupancy monitoring and designing occupancy-based HVAC systems, though typically the sensors are able to detect the presence and not the number of occupants in real time. Consequently, ventilation is generalized based on the maximum occupancy and not the real number of occupants leading to energy consumption more than necessary. In offices and places with dynamic occupancy patterns and constant rate of ventilation, usually, the zones are colder than necessary due to injection of a high volume of fresh outdoor air. This will push occupants to use heaters even on warm days [9].

CO_2 buildup has been used typically in HVAC systems as an indicator of the number of occupants [10, 11]. Warren et al. employed CO_2 sensors to estimate occupancy and their simulation results indicated 53% reduction in heating energy consumption [12]. Although CO_2 meters are being used widely for controlling HVAC systems, from the moment that occupants enter a zone until they generate enough CO_2 to be detected by the sensor, it takes 10 to 20 minutes. Hence, another sensor is needed to identify the number of occupants in real time.

The typical binary presence sensors such as Passive Infrared (PIR) and Ultrasonic may be the solution for the lighting system. However, in the case of HVAC systems, the real-time number of occupants is required. PIR sensors were employed by Agarwal et al. to detect movements in each room and they decided to only heat the rooms which were detected occupied. This sensor was prone to the false OFFs due to nearly motionless occupants [13].

In order to leverage more accurate sensors in terms of occupancy, Erickson et al. utilized network of cameras to monitor occupancy for HVAC consumption. The accuracy of the system was more than 80% and they reported the HVAC energy consumption could be decreased by 30% to 40%. However, cameras are so expensive and it could raise privacy concerns [5].

Li et al. used RFID tags to detect and track occupants to calculate the heating demand for thermal zones accordingly. Their monitoring system achieved an accuracy of 88% and 62% for stationary and mobile occupants respectively [14]. In another study, RFID was used to detect occupants but their system only spotted the ones passing doors or at the tables [15].

In spite of the fact that these methods are dedicated to monitor occupants and lower the energy consumption, they need extra costs of installation and maintenance.

Cerpa et al.[16] estimated the cost of $147000 for just the manufacturing and deploying camera and PIR nodes in a 3 story building. Balaji et al. [17] reported the expense of $120000 for simple motion sensors for their 5-floor building.

In order to cut the expense of extra sensor installation and due to widespread utilization of WiFi technology, use of WiFi data has gained more interest in recent years. WiFi data has been widely used for indoor localization [18], Crowd density estimation [19], social interaction monitoring [20] and occupancy monitoring [21]. Melfi et al. leveraged existing wifi technology to monitor occupancy and manage plugged-in devices i.e. desktops. The idea was to use the number of users connected to the internet as an indication of number of occupants. They found the electricity usage and the number of WiFi hosts had the same pattern. The advantage of this method is that there is no need to install extra sensors as most of the commercial buildings are equipped with WiFi technology [22]. Balaji et al. used WiFi connection of the smartphones and the device MAC addresses to identify the owner of the single occupant offices. They marked the MAC address which was mostly near the office, as the owner of it and ran the office heating systems only when that MAC address was present in the WiFi system [17]. In SocialProbe, probe requests, which are being sent between smartphones and access points to manage connectivity was used to monitor the interaction of the occupants [20].

2.1 Occupancy matters

Despite all the investments in new technologies and physical enhancements, the energy conservation goal is not fully achievable without taking into account the role of occupant [23].

Since all these technologies are being operated by humans, Misstep of the human contribution can result in the collapse of the whole energy saving mission [24]. Occupancy behavior refers to the movement, habits, and interaction of the occupants with the building elements to satisfy their needs which affect the energy performance of the building [25].

Occupant behavior is the main factor that results in huge gaps between real and predicted energy consumption of buildings [26]. The behavior of the people in the building consist of two parts: 1) when they are in the building and how many occupants are in each zone. 2) How they are interacting with elements of the building i.e. windows, blinds, AC, and appliances [27]. If a building operates based on a real time occupant population instead of a fixed schedule, the energy saving could be quite a lot.

Because of the stochastic characteristic of the human, energy consumption does not only vary from one person to another, it tends to change in time and space. It is hard to predict and model the energy consumption of the buildings, especially the ones which tenants have the right and means to adjust the devices based on their needs [22]. To better understand the difficulty of predicting occupant behavior on energy usage Li et al. measured the air conditioning electricity use of 25 residential unit in the same building in Beijing. Although all the households had the same envelope and were located in the same climate, the electricity consumption varied significantly (Figure 4). Therefore, in this case, human behavior is the driving factor of energy consumption rather than the building design.

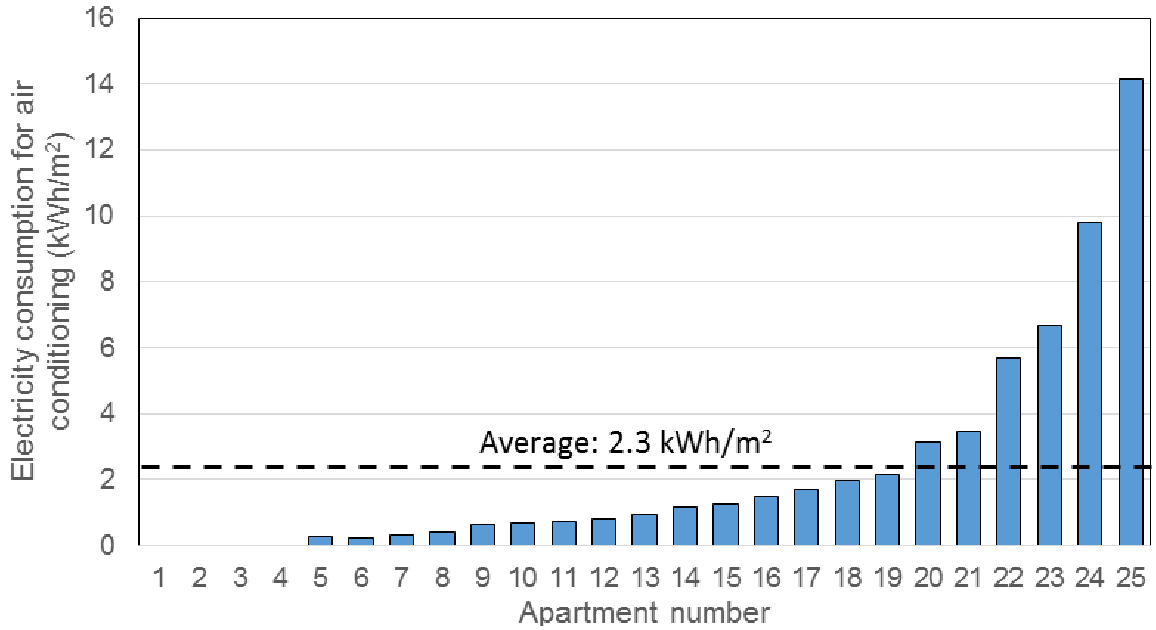

Figure 4- Air-conditioning electricity use of 25 households in Beijing [22]

Occupant behavior is affected by different factors. Warren and Parkins indicate that the possible factor that makes tenants open the windows in winter is the air quality inside the room, while in summer the outside noise is the reason why they close the windows [28]. Behavior can even be influenced by physiological parameters. A study in naturally ventilated office buildings indicated that occupants who tend to enjoy the view and connects to the outside used to keep open the blinds [29]. On the other hand, the ones who were facing other buildings tends to close the blinds due to privacy issues [30].

There are three type of energy-related behavior: ***1) conservation behavior***: the ones who mind the amount of energy that is being used and tries to save the energy, ***2) average behavior***: who don't try to save energy, but also don't desire to waste it either. ***3) Wasteful behavior***: not only attempt to save the energy, but their action could even lead to energy wasting. According to Nguyen, "occupants with careless behavior can add one-third to a building's designed energy performance, while occupants with conservation behavior can save a third"[25].

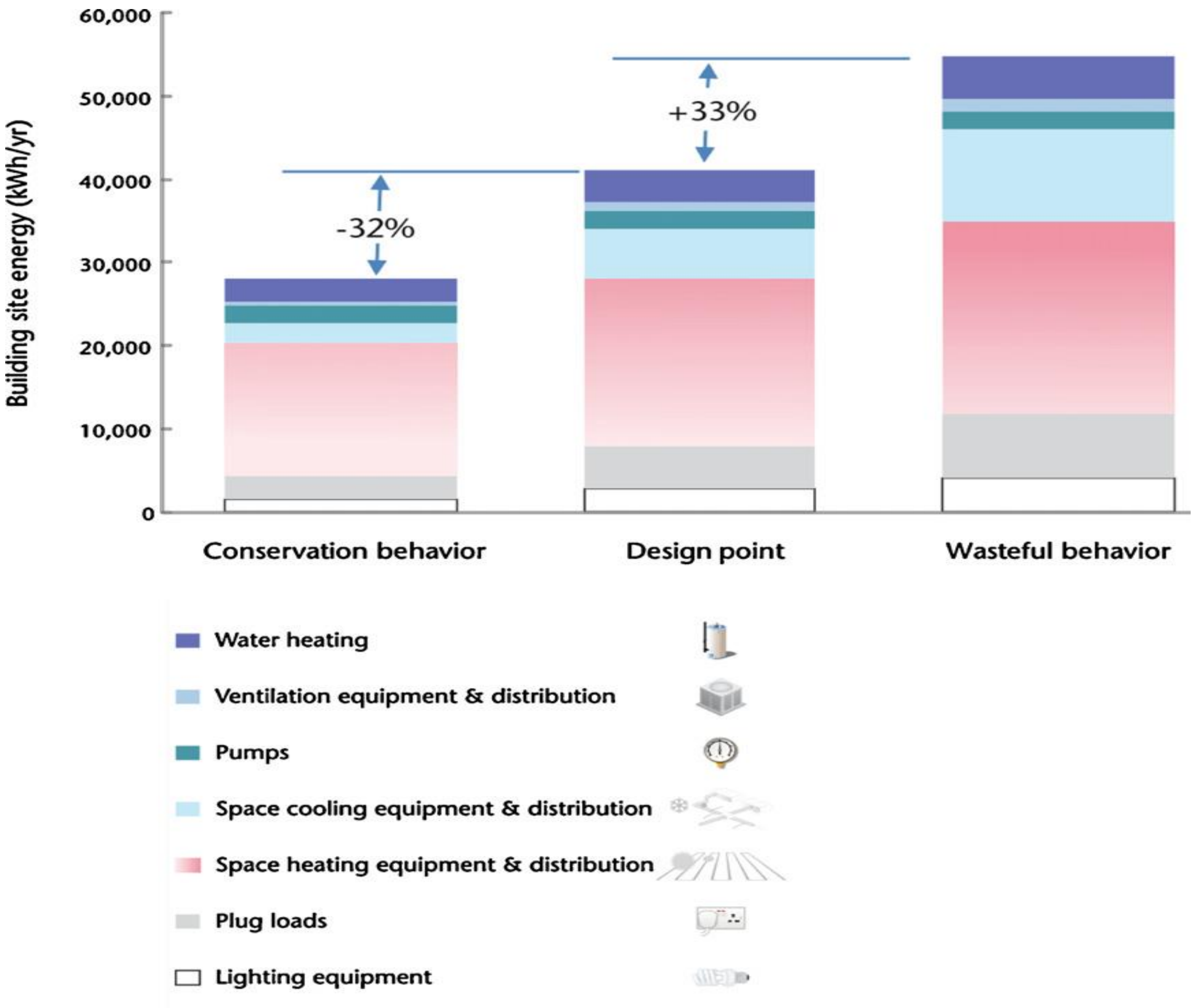

Figure 5- Energy consumption based on occupants with different behavior [25]

Ouyang and Hokao reported a survey on the potential of energy saving by raising awareness on energy-conscious behavior in residential sectors. 124 Households were divided into two groups. One was taught to improve their behavior towards energy saving and the other one was asked to remain unchanged. The electricity bills showed that educating occupants to promote their energy-related activities could decrease consumption by more than 10% [31].

There are multiple approaches to lessen energy consumption through technologies and using renewable energies, but these solutions are costly in terms of capital expenses and installation and maintenance costs. It could also disrupt occupants comfort and takes too much time to be deployed. Whereas, Behavior change can be implemented without spending too much of money and can be accomplished quickly. Energy saving opportunities due to considering occupancy behavior is often ignored in defiance of the fact that it is referred to as significant as the technological solutions [32]. Furthermore, investment in the energy related strategies may take a long time to be recouped and organizations are likely to place great emphasis on the initial costs rather than life-cycle cost.

In summary, change of behavioral pattern has a huge opportunity in energy saving compared to technological solutions. It has zero cost, no technical knowledge and it can be implemented easily in both future and existing buildings.

2.2 Energy Intelligent Buildings

Energy intelligent building refers to a building with the ability to monitor its occupants and technologies to optimize control of appliances with the aim of saving energy [25]. Different sensors and approaches had been used to monitor occupancy in order to save energy.

The university of California presented SPOTLIGHT [33]. The approach is that occupant uses RFID tag which is supposed to detect the distance of the user and building appliances (i.e. TV, lights, PC) this information report the amount of wasted and used energy.

According to Sensible Building Science which is a University of British Columbia based startup, using occupancy pattern to control the university's building management system could lead to 7% energy saving annually.

Mozer used Neural Network to control the lighting system of a building and it was able to gain a remarkable electricity reduction but it was along with occupants' discomfort at some levels [34].

The BODE project used camera network solution to develop a methodology that predicts user presence in the building and keep track of user movement in different zones of the building (Bode Project 2011).

ARIMA has been developed by National Research Council Canada to predict the building demand. They used wireless sensors in Canada comprising laboratories and individual work spaces to collect the data.

2.3 HVAC systems

Air quality is a critical factor that affects the productivity, health and comfort of the building occupants. In addition to adjusting the temperature according to the weather condition, occupants also need fresh air. This fresh air comes from outside and after adjusting to the desired temperature it is replaced with the indoor air that may contain CO_2 and other pollutants. Recirculated conditioned air in heavily insulated buildings may carry contaminant that will lead to health problems. Thus, air filters must be used to remove dust and pollutants. It is HVAC system's responsibility to achieve all these tasks.

To better understand the Heating, Ventilating and Air Conditioning system, the procedure of a simple system is explained. Chillers and heaters produce hot and cold water. Air handler unit brings in external air and heats or chills it with hot or cold water until it reaches the desired set point. Fans then spread the conditioned air in the building. As the difference between the air supply and set point increases, the system uses more energy to supply conditioned air.

A typical design for a ventilation system in most buildings is to set it for maximum occupancy or area capacity. However, there could be less occupant and even no occupant for some periods. Therefore, some rooms are being ventilated needlessly. Ventilation rates in recent buildings are set to ASHRAE standards, however, in older ones, there are chances that they are being either over ventilated or under ventilated at some cases.

Demand control ventilation (DCV) is denoted for altering the airflow to reflect changes in area occupancy (ASHRAE 2013). The amount of fresh air injected to each zone is controlled by "dampers" that are installed per zone. Places with continues variation in occupancies such as gym and cinema benefit more from DCV. In the U.S the opportunity to decrease heating and cooling loads by using DCV has been estimated almost 20% annually (ASHRAE 2013).

Figure 6 illustrates the application of CO2 based DCV. In this method, ventilation will receive a signal from CO2 meter and then ventilates each zone proportion to the CO2 concentration. Another alternative to CO_2 could be turnstiles that show how many had passed it.

The CO_2 concentration is currently being used to measure occupancy in most of the HVAC systems. However, CO_2 level is unlikely to change in bigger areas with few occupants. This method has more flaws that will be discussed in next sections.

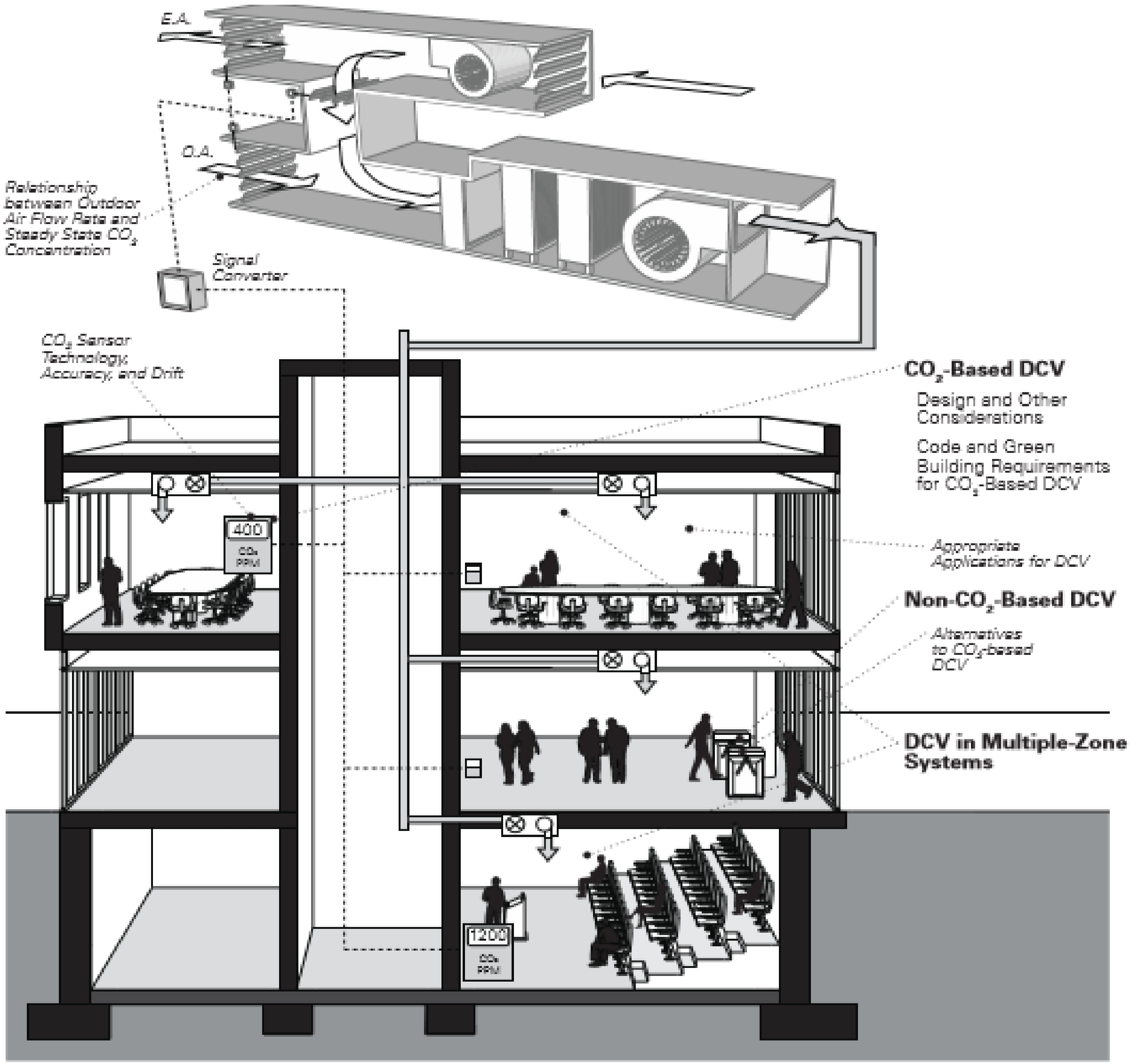

Figure 6. Demand Control Ventilation (ASHRAE Indoor Air Quality Guide)

2.4 Indoor localization

Global Positioning System (GPS) is widely being used for outdoor localization but this system fails when it comes to indoor localization as it loses its coverage. [36]. The following indoor localization techniques had been studied in literature:

1. Triangulation:

   In this method, the position of the device is estimated by leveraging the angels of the arrival of the received signal of the device from reference points at some stations.

2. Proximity:

   If a sensor can detect a presence it means that subject is in the vicinity of that sensor. But this location is not specific and it's related to the sensor coverage. The strongest signal will be marked as the location.

3. Fingerprinting:

   This method uses RSSI database and has two phase, before localization it generates a map at some known points by collecting the received signals at each reference point. This is called offline phase and it is almost the same in most research. It then compares the signal at unknown points with the map and the closest match will be marked as the location of the device. The second phase is called online phase. Different statistical and machine learning methods such as Neural Network [37] Decision tree [18] had been used to localize the device.

   From all these methods fingerprinting has been reported as the most precise one [36] although using RSSI for localization has some drawbacks. The RSSI value varies with time and the indoor characteristic i.e. structure of the building, furniture and people movements have an effect on the received signal. [38]. Therefore, a big database is needed to overcome the mentioned problems. Some authors have proposed to use external sensors and resources like RFID and landmarks (Chen et al. 2015; Chen et al. 2005) but this will add to the cost of hardware and computation. Moreover, it has been observed that RSSI value varies with the change in the direction of the device [41]. This will need a huge database for different orientations at each reference point. With all challenges, fingerprinting method is still the most sought after approach for indoor localization as it reported high accuracy and low cost compared to other methods. This method will be discussed in more detail in upcoming chapters as it has been selected for this thesis.

2.5 Occupancy Monitoring Approaches

Different methods and sensors had been used in published papers to detect indoor activity i.e. presence and number of occupants and localization. Some of these methods are summarized as follow:

2.5.1 *Passive Infrared (PIR)*

PIR sensors are the common approach of the many projects since they are cheap and easy to install compared to most of the sensors and most importantly they are not violating occupants' privacy. Passive Infrared (PIR) sensors come in a pair and detect the direction of the movement in their range. They are mostly being used in efficient buildings for lighting systems. It has a 0 or 1 output that shows if somebody is present or not in a particular zone. When a warm body passes by, the heat image would change. Therefore PIR is able to detect movement but sometimes if there is a nearly motionless occupant, PIR is not able to detect whether space is occupied or not. Moreover, if a few people pass PIR sensor at the same time it may not be able to count correctly. It is also possible that it detects non-human movements or got triggered by heat current released from the heating systems.

Despite all the drawbacks of PIR Maniccia et al. showed that if PIR sensors are being used in small offices with a time delay setting, it can actually save energy up to 43 percent. It also showed that the office is occupied for only 30% of the time [42]. A more recent study carried out by Goyal et al. in a small office with 3 occupants and PIR sensors. The HVAC was operated only when PIR detects presence. It resulted in 40% reduction in energy consumption [43].

SPOTLIGHT used proximity sensors to detect the occupants' interaction with building appliances. (Y. Kim 2008). Newsham and Birt developed ARIMA model using door closure sensors and PIR motion detection along with the gathered information of the number of people logged into the network, the outdoor temperature of a building in Ontario to estimate power demand [44]. Hagras et al. used pressure pad sensors to determine if somebody is sitting on the chair or a bed. He also used a custom code that would detect the network based activities on the computer (H. Hagras 2004).

2.5.2 *ULTRASONIC*

Ultrasonic sensors are reliable devices that are widely being used for object detection or level measurement. They are suitable even for toughest conditions. The surface of the sensor is always dirt-free as it is constantly vibrating. They have the ability to both sense and transmit. These sensors have the ability to detect the presence and unlike PIR the object must not be in sensor's line of sight since ultrasonic waves could cover all the space. As the sensor is so sensitive to movement, it is more prone to false detection. Similar to PIR, the ultrasonic detection system is not capable of counting the number of occupants. Therefore, it is commonly being used in the lighting system that just needs binary presence detection.

Floyed et al. used ultrasonic for the lighting system of an office and a school and observed an increase in energy consumption. Later he found out that the rise in consumption was due to false triggering of the ultrasonic sensors (i.e. paper printing, occupants stopping by a room for a few seconds) [45].

2.5.3 *CAMERAS*

Cameras have the ability to both detect occupancy and count the number of people available in the zone [46].

Mahdavi et al. reported a survey on 48 offices in which they recorded occupants with cameras and found out office workers are away from their workstations more than 50% of the time, but lights and equipment remain up the whole time [47].

Varick et al. developed SCOPES, a network of cameras to detect occupancy. They reported 20% energy saving while maintaining ASHRAE standards [5].

The problem with cameras is that they are expensive and interferes with occupant's privacy and comfort [48]. In addition to that, lighting quality can influence the sensor performance. Multiple people passing at the same time may also lead into under counting.

2.5.4 *RFID*

Radio Frequency Identification (RFID) consist of some tags that are carried by the occupants and some readers. This sensor is capable of detecting and counting the occupants. As it uses RSS (Received Signal Strength) it can also detect the location of the badges in room-level accuracy, but only the occupants with a tag can be detected.Moreover, since each badge is assigned to one person it can violate the occupants' privacy. Rahman et al. used RFID tags on a high rise building to detect the movement of occupants for evacuation situations. The result showed near 100% accuracy in locating [49].

Lee et al. used RFID tags to detect users but the system could only detect the ones who were sitting at tables or passing doors [15].

Although these sensors are suitable for conditions that location must be accurate, this accuracy is along with a high cost as it needs one tag for each occupant.

2.5.5 *Carbon dioxide meter*

As the amount of CO_2 will change by the number of people present in the room, the CO_2 meter is being used as a detector of occupant presence. This sensor is not capable to count the number of occupants but it's able to detect the increase in CO_2 level. That is the reason that for ventilation systems CO_2 meters are being used to control the fresh air injection at most buildings for almost 20 years [11].

Warren and Harper reported using a CO_2 controlled ventilation system can save 50% energy in an auditorium that uses 100% fresh air [12].

One drawback is the point that it takes a while from the moment people get inside the room until they produce enough CO_2 to be detected by the sensor. Also, other factors such as the wind, the location of the sensor and occupants activity can interfere with the sensor task. Therefore, this sensor is not a good indicator of the people presence.

2.5.6 *WiFi monitoring systems*

This method makes use of a network of WIFI sensor nodes to detect wifi equipped devices i.e. smartphones. WiFi probe requests are sent from the WiFi enabled devices like smartphones or laptops to scan a specific network or all the available networks. The device will broadcast a "who is there" packet, which consists of the device MAC address and RSSI. Media Access Control is a globally unique number that has been assigned to each computer hardware. These nodes can sniff these packets or trigger the devices to send probe requests. This method can be used for marketing purposes, for example in shopping centers. As each node has a range, with analyzing the data it can be understood which shops are visited frequently by costumers or where in the mall they spend more time. Moreover, this method had been studied for search and rescue purposes. In a study a WiFi enabled drone flown over the search area and triggered the user smartphone to send probe requests [50]. Verbree used a commercial Wi-Fi monitoring system to count the visitors in the museum and discovered some major drawbacks of this system. They found out that scan frequency is longer than detecting the occupants who were moving from one room to another room. Also, it was hard to detect the phones that were in sleep mode (no active) as the ping request was sending in a longer interval [51]. Christensen et al. made use of existed access points ARP (Address Resolution Protocol) table to detect the occupancy presence in an office. (Christensen 2014). MAC address is stored in the ARP table of the routers as long as the hosts and nodes are exchanging packets. If there are no more packets forwarding between host and AP for a predefined timeout duration, the MAC entry will be removed from the table.

Wifi sensors can show the number of occupants in a zone, though there are some MAC addresses that are always on the table, which belongs to the fixed wifi equipment.

Compared to other technologies Wifi sensors are much cheaper and it does not need a clear line of sight as it can pass the walls and cover a bigger area. It would be completely free of charges if you can access the already installed Access points.

The drawback of this method is that not all the occupants carry just one device. For example, they may use their smartphone and laptop at the same time, so there will be two MAC addresses in the table, though there is just one person in the zone [22].

If the number of occupants is limited and they are local, by assigning each MAC address to its own occupant we can estimate the total number of occupants.

Leveraging Wi-Fi sensor networks could also provide data related to area utilization. This data will show how often and how many occupants are actually using the provided space. It will help in constructing more cost-efficient buildings.

2.5.7 *WiFi based localization systems*

These systems use the signal strength receiving from the wifi enabled devices to locate the device. The moment MAC address is detected by the wifi sensor network, the RSSI (Received Signal Strength Indicator) associated with each Mac address is also captured. The value indicates how well the device is receiving a signal from the access point. Dependent on the manufacture it ranges between 0 decibels (amazing) and -110 decibels (very poor). As a device is getting closer to the AP the RSSI value will increase. So it can be used to estimate the distance of the occupant from the AP [52]. RSSI also can be used for triangulation [53]. If there are three or more sensors with three different RSSI value of one device, the location of the device could be measured using triangulation and finger printing method [52]. There are three ways to estimate the position using RSSI value:

4. Triangulation:
   As the device gets closer to the AP the RSSI value will increase. In this method, the position of the device is estimated by calculating the distance of three AP with known positions. Due to the unpredictable noise level, this method is not so reliable.
5. Proximity:

The strongest RSSI from the AP will be marked as the area in which the occupant must be present. This method is not specific and the device can be anywhere in the area covered by the AP.

6. Fingerprinting:

   This method compares the RSSI received in online mode with the database in which RSSI had been collected at some reference points. This method will be explained more as it is the approach that has been used in this thesis.

Khoury and Kamat in an indoor positioning project investigate the feasibility of three different methods: wireless local area network (WLAN), indoor global position system (IGPS) and RFID system. They used RSSI (received signal strength) to estimate the location of occupants. The result from IGPS and RFID was more promising than the WLAN (more than 2-meter error). In terms of installation and cost, WiFi system does not require much. Moreover, this method does not need a clear line of sight and can actually go beyond the walls [54]. Another advantage of wifi detection based sensor is the overlapping coverage, which makes it easier to track down occupants.

2.5.8 *Summary of monitoring sensors*

In summary, Ultrasonic sensors are expensive with high sensitivity which may not be good to sense the presence, especially in open floor plans. It also is not capable to count the number of people. PIR detection systems are good for controlling the lighting system that just needs presence detection. However, if there is an occupant who is not moving that much, the sensor may detect no presence in the place. RFID and cameras may raise privacy concerns as the occupants' identification are exposed. CO_2 sensors need to be installed in the breathing zone to detect correctly, as it's not applicable they will be installed on the ceiling. Also from the moment that occupant enters until the time that sensor can actually detect the CO_2 level it may take half an hour or even more. In large spaces, if one person enters with a smartphone connected to wifi, his Mac address will be added to the number of Mac detections. However, in the case of CO_2 sensor, it may not detect any difference in the CO_2 amount. On the other hand, using already installed Access points needs no additional cost or application. It can detect both presence and location of the occupants. This sensor just detects the MAC address and possibly the estimated location of the WIFI-enabled device. The MAC address could be stored in hash version. Therefore, it does not raise privacy issues compared to cameras. Another shortcoming of most of the explicit sensors is their lack of networking to transfer data for collection and post experiments.

2.6 Artificial Neural Network

Artificial Neural Network refers to a human brain inspired programming model which allows computers to learn from observational data. Neurons in our brain receive multiple messages made by our actions and senses. This inputs are shared by other neurons and then passed to the brain to take action. Experience and new learnings help neurons to modify themselves regularly.

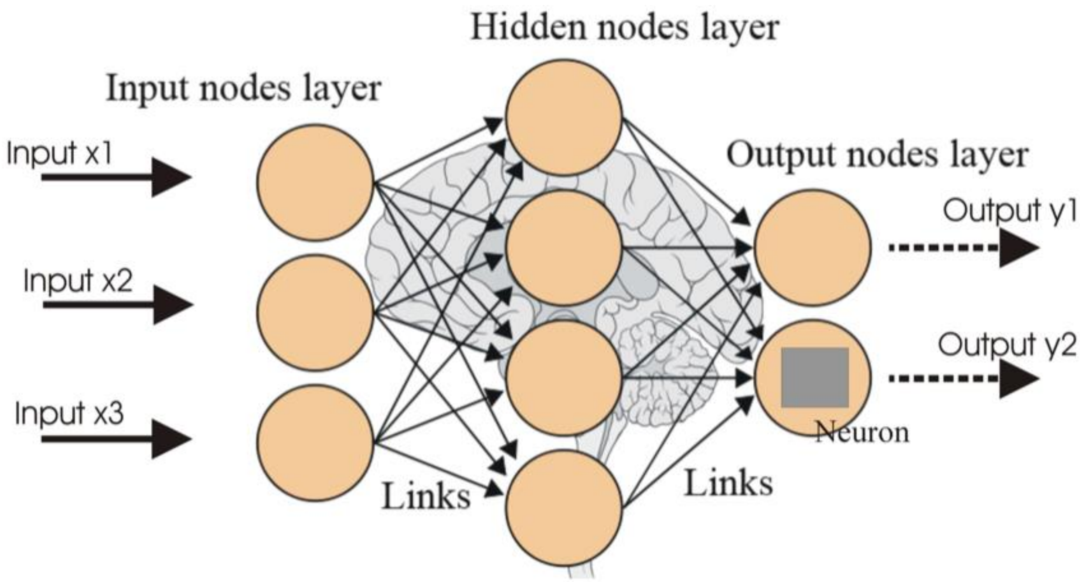

Figure 7. A simple neural network design.

Despite the traditional programming, neural network figures out the solution to complex problems by observing the patterns in a given data.

A typical neural network model is shown as a number of layers contains nodes: An input layer, an output layer and a number of layers in between known as hidden layer/s (Figure 7). Each neuron can have multiple inputs (x_i) with an associated weight (w_i) and an activation function (f) that applies to the weighted sum of the inputs (Figure 8).

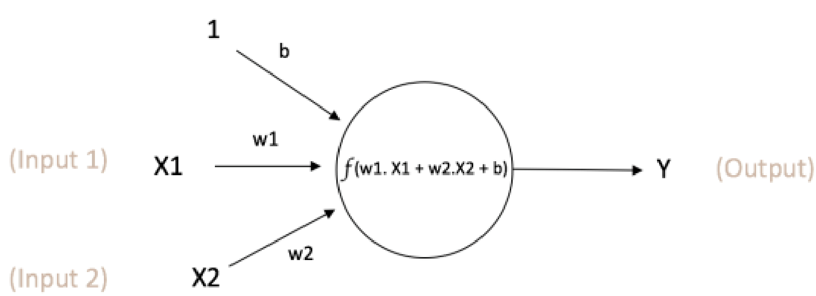

Figure 8. A single neuron of Neural Network model.

One of the most frequently used activation functions is the sigmoid function:

$$f(x) = \frac{1}{1+e^{-x}} \quad (1)$$

In neural network method, neurons change their weight through a training algorithm and evolve continuously through an optimization algorithm (e.g. gradient decent back propagation algorithm) until the cost function is minimized. The cost function is simply defined as the summation of the squared error between the desired target and output. General formulation of a simple case of one hidden layer neural network model is as follow:

$$Z^{(2)} = XW^{(1)} \quad (2)$$

$$a^{(2)} = f(Z^{(2)}) \quad (3)$$

$$Z^{(3)} = a^{(2)}W^{(2)} \quad (4)$$

$$\hat{y} = f(Z^{(3)}) \quad (5)$$

Where X is the input matrix, $a^{(2)}$ is the activation function of the layer and $\hat{y}$ is the calculated output value. Z is also defined as the sum of the matrix of the weighted inputs (Figure 9).

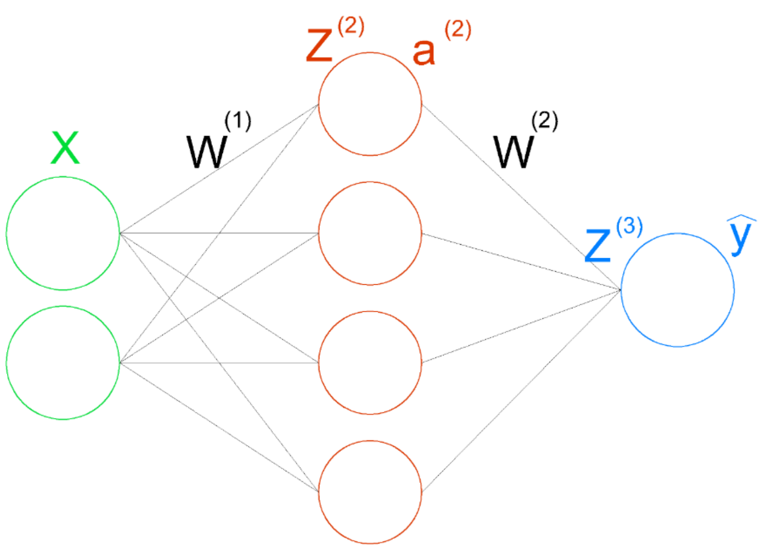

Figure 9- definition of the aforementioned one hidden layer neural network

Currently, the neural network is being used in understanding and solving complex problems such as image processing, speech recognition and complex market predictions.

For instance, Problems such as handwritten letters recognition would be so hard to solve with traditional programming. However, it can be solved more efficiently with NN. The network will take a large example of handwritten letters for training purposes and then develop some rules for each letter. The next time it sees a “b” letter, it automatically detects it based on previous learnings. As the input data size increases, NN will generate more rules and learns more about handwritten letters. The network evolves continuously until it reaches a state where it cannot improve anymore. The accuracy of the model is evaluated by how close the output is to the desired target.

Neural network algorithms are being categorized in two major subcategories: Supervised and unsupervised. The supervised method uses a training data set, which contains a set of input and the associated response to it. Therefore, the network builds a model based on the training dataset and make predictions for the new dataset. Classification (discreet output values) and Regression (continuous value outputs) fall into this category.

Unsupervised method uses only input data, without knowing the output values and finds hidden pattern and similarities in the input data, clustering is an example of this method (Goodfellow et al. 2016)

Chapter 3. Experiment Setup and Methodology

This thesis consists of two major parts. The first section is allocated to designing an occupancy based ventilation system, the second section is related to mapping occupant location to an even more optimized ventilation system.

3.1 Occupant-based ventilation system

Heating ventilation air conditioning (HVAC) systems are the main focus of nowadays researches as it is the main energy consumer of the buildings. The existed buildings are using a predetermined HVAC system without flexibility which is mostly based on a fixed schedule related to the start of the day, end of the day, temperature and humidity inputs. These inputs are insufficient as there may be a room without any occupant in it but it is being heated or cooled like the other zones of the building. The ventilation system is fixed based on the room capacity without taking into account the real number of occupants. For example, lecture halls have the capacity of 100 people but only 20 may attend. A large number of researches have proposed using Camera, CO2 meter, Motion detector, RFID tags, PIR sensors etc. to estimate the occupancy. But these sensors require extra money for installation and maintenance, especially at large buildings. Nowadays most of the buildings are equipped with WiFi access points in order to provide internet access to its occupants and it has attracted the attention of researchers to use this infrastructure in order to control light, HVAC and so on.

Media Access Control (MAC address) is a universally unique ID that will refer to a network device. The first 3 bytes is called OUI which defines the ID of the manufacturer and the last 3 bytes is the serial number for that particular device. Since this address is unique per device universally it can also be used to monitor the person who is using it. Melfi et al. used the number of released IP addresses to estimate the number of occupants [56]. The ip address is layer 3 addresses that will assign to devices by network administrators. Melfi had claimed that as one occupant may have more than one WiFi equipped device, the number of released addresses are not exactly equal to a number of occupants but the pattern is somehow similar to electricity usage pattern.

We used the 7th floor of Ryerson YNG building for our experiment. This building houses shared research labs for graduate students of Ryerson and some other labs that are not related to Ryerson students, those labs had been excluded from our zone of interest. The elevator and stair way areas have also been discarded (Figure 10). These areas are usually not ventilated as other spaces of a building. There are 5 Ryerson access points on this floor. We assume if one access point detects the user MAC, thus this user is at the range of the AP and in the building. A survey was conducted to find out how many WiFi devices the students use in the building. As expected, all of them belongs at least one smartphone and since they have access to the desktop computers in their office, they were unlikely to use their personal laptops. Figure 10 illustrates our zone of interest. Suite 703 and 704 were excluded as they are not part of Ryerson property. Therefore no one in these suites has access to Ryerson network.

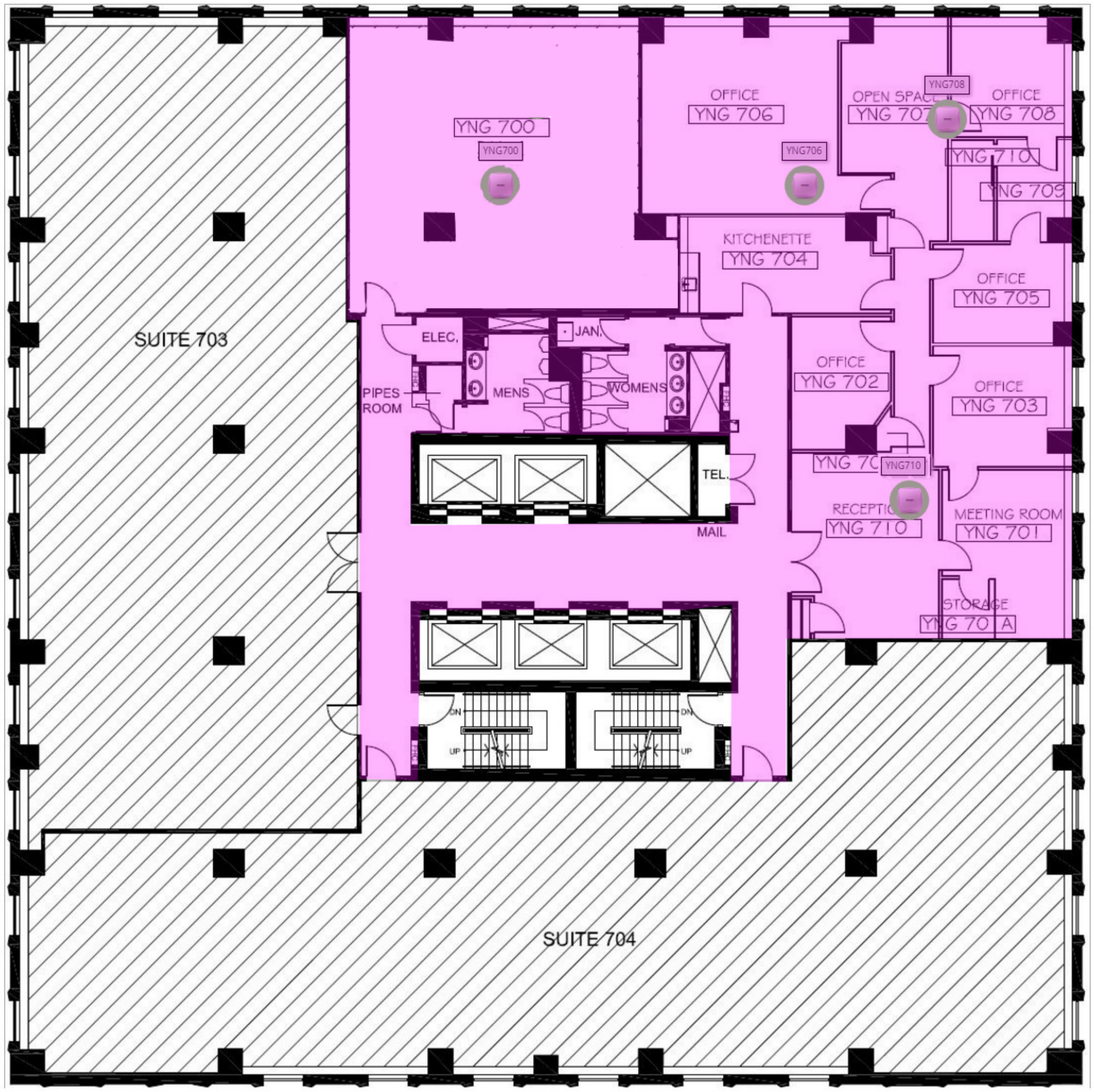

Figure 10. YNG 7th floor plan zone of interest.

It is possible that when an occupant is moving in and out of his office he does not carry his phone, he may leave his phone at home or the phone runs out of battery. Each system that uses WIFI data as an indicator for presence detection struggles with this problem [17]. Thus, since we are using the MAC address to indicate the presence of the clients, we assume wherever the phone is, the occupant is most likely there and the phone is always connected to the internet.

The MAC address data of users connected to YNG 7th floor APs has been provided by Computing and Communicating Services (CCS) staff of RYERSON for 5 minutes intervals and for 3 weeks. The format contains the date, time and the mac address along with the RSSI associated with each MAC address (Figure 11).

```
2017-02-19-08:51:01 80:00:0b:c3:74:db 71
2017-02-19-08:51:01 3c:e0:72:a9:bb:26 70
2017-02-19-08:51:01 f4:8c:50:98:01:f2 0
2017-02-19-08:51:01 e8:4e:06:35:18:dc 76
2017-02-19-08:51:01 ac:2b:6e:ba:13:0f 77
2017-02-19-08:51:01 c0:ee:fb:26:39:c7 77

2017-02-19-08:56:01 04:54:53:07:e4:33 71
2017-02-19-08:56:01 80:00:0b:c3:74:db 73
2017-02-19-08:56:01 3c:e0:72:a9:bb:26 70
2017-02-19-08:56:01 da:a1:19:1a:fd:0f 60
2017-02-19-08:56:01 f4:8c:50:98:01:f2 0
2017-02-19-08:56:01 cc:20:e8:64:2c:df 74
2017-02-19-08:56:01 64:a3:cb:02:33:2b 71
2017-02-19-08:56:01 e8:4e:06:35:18:dc 76
2017-02-19-08:56:01 80:00:0b:c3:74:db 64
2017-02-19-08:56:01 96:b8:9f:a4:f0:3c 72
2017-02-19-08:56:01 80:00:0b:6f:64:0e 70
```

Figure 11-Raw data of YNG building

Here in this thesis, we have used MATLAB 2017 to code the corresponding calculations. Using MATLAB code we have packed all the data from all the five different access points in one data file and restructured them accordingly. we calculated the unique times and then the number of unique MAC addresses in each unique time interval has been counted. Doing this we were able to plot the total number of clients in 5 minutes interval which is our sampling time.

Figure 12 shows the number of MAC addresses for the 3 weeks of data collecting at YNG 7th floor. As it is clear, the number of MAC addresses fluctuates within each day. Moreover, there is a significant difference between these 21 days in terms of the number of unique MAC addresses. The maximum number of MAC addresses is 25 on February 19th, whereas on March 5th and March 9th the maximum is 67 and 13 respectively. This will clearly represent using a fixed pre scheduled ventilation is not an optimized solution and results in too much waste of energy. Moreover, March 5th shows a sudden rise of MAC addresses that are clearly much more than the maximum occupancy in previous days and a fixed ventilation would lead to an under-ventilated situation.

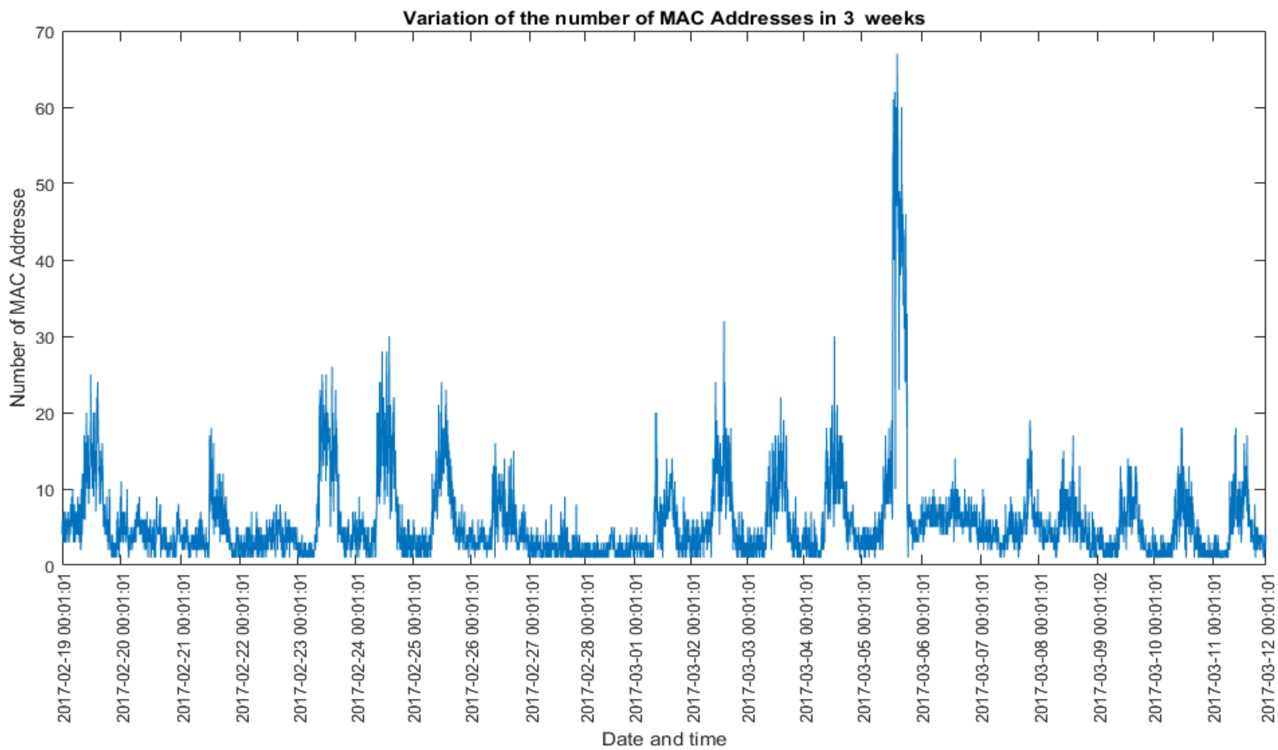

Figure 12. Number of MAC addresses in three weeks.

The below figure shows how users MAC addresses change through 1 day. Around 8 P.M we observed a rise which is the start of the working day. Between 12 P.M and 2 P.M, we observed higher fluctuation in number of MAC addresses which could be due to the lunch break. After 4 P.M the number reduced again which could be the result of ending of the working hours. This figure shows the number of occupants is changing continuously and a fixed airflow is not an optimized solution for this space (Figure 13).

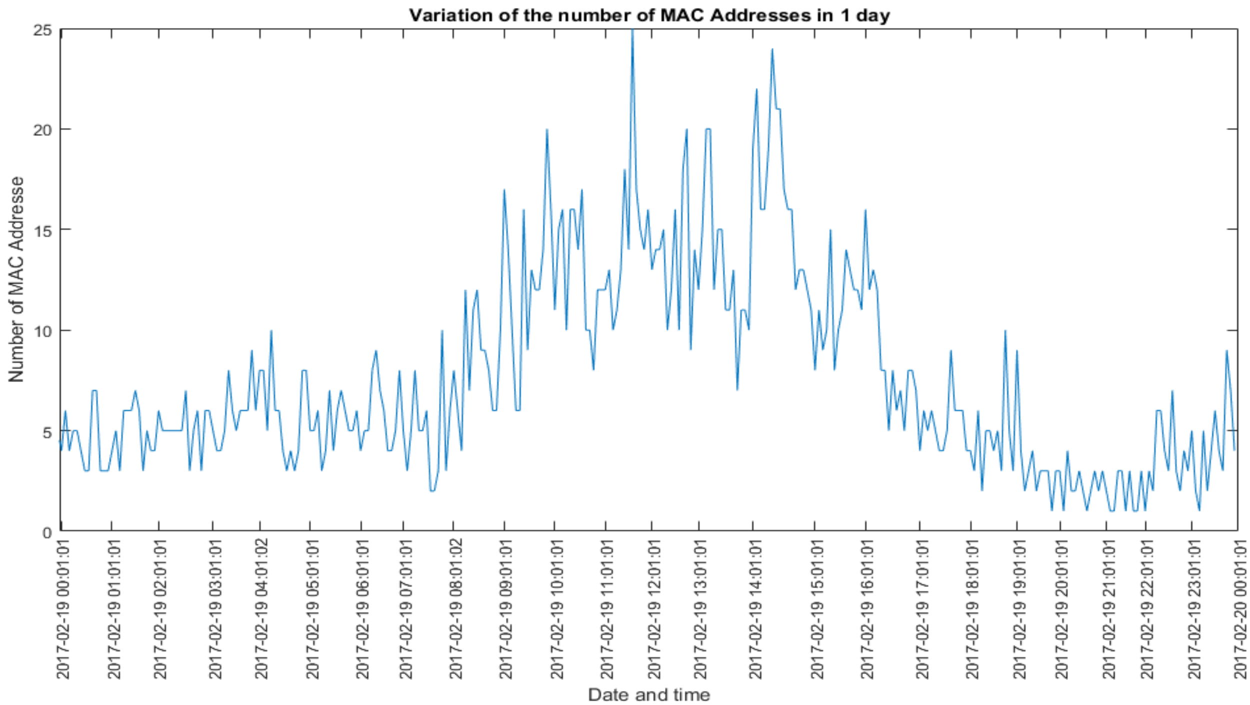

Figure 13. MAC address variation in one day

As illustrated there are some MAC addresses that are always present even during the night. These MAC addresses are most probably some WiFi devices that are not related to occupant presence. These ones are not having the attitude of an occupant i.e. coming and going or moving from one place to another. We decided not to discard these MAC addresses and leave them as is. This acts as a safety factor for those who forgot to bring their phone, or the ones that had lost connection to the internet. Moreover, as we wanted to compare the new design of ventilation with the old one (working from 7 A.M – 7 P.M for maximum occupancy), the after hour present MAC addresses are not important and will not lead to wasting the energy.

3.1.1 *Evaluation*

To evaluate this method we observed the number of occupants at different random times and compared the results. We count the number of people who were entering and exiting the 7th floor for 10 minutes at each time and then compared our observation with the records of Unique MAC Addresses. Table 1 shows our result.

Table 1 System Evaluation

#Observation	Start time	End time	#Recorded Mac at start	#Recorded Mac at end	#occupants entering	#occupants exiting	Change in number from observation	Change in number from recorded data	Error
1	'2017-02-19-09:41:01'	'2017-02-19-09:51:02'	12	20	12	4	8	8	0
2	'2017-02-19-10:06:01'	'2017-02-19-10:16:01'	15	10	0	4	-4	-5	1
3	'2017-02-19-14:01:01'	'2017-02-19-14:11:01'	19	16	0	2	-2	-3	1
4	'2017-02-23-15:41:01'	'2017-02-23-15:51:01'	14	12	0	2	-2	-2	0
5	'2017-02-23-09:21:01'	'2017-02-23-09:31:01'	16	23	13	3	10	7	3
6	'2017-02-23-11:01:02'	'2017-02-23-11:11:01'	21	15	1	6	-5	-6	1
7	'2017-03-3-11:21:01'	'2017-03-3-11:26:02'	19	19	2	2	0	0	0
8	'2017-03-3-16:41:01'	'2017-03-3-16:51:01'	9	12	5	2	3	1	2

Error is defined as follow:

$$E = (Observational\ Data - Recorded\ Data) \tag{6}$$

As illustrated, the observation is so close to the recorded data, in most of the times, there is no difference between the change of numbers in observation and in recorded MAC address data. The difference could be because of the fact that some may not have connected to the Wi-Fi or some may have laptop and smartphone connected to WiFi at the same time. But these differences are not huge and could be ignored. Moreover, accurate data is crucial in many applications, however, Ventilation is less sensitive in this regard (Christensen 2014) .

3.1.2 *WiFi-based ventilation system in comparison with traditional ventilation System.*

A typical design for a ventilation system in most of the building is to set it for maximum occupancy. However, there could be less occupant and even no occupant for some periods. In recent changes to ASHRAE standard, ventilation for unoccupied zones can be turned off [57]. For instance, classrooms ventilation could be set proportion to the number of students, then turn down to a minimum between classes and shut down overnight. It must be turned on next day prior to the class starting time for an expected number of students [8].

Demand control ventilation (DCV) is denoted for altering the outdoor air flow rate to reflect changes in area occupancy (ASHRAE 2013). The CO_2 concentration sensors are being used recently to measure occupancy. However, CO_2 level is unlikely to change in bigger areas with few occupants. Moreover, it takes some time from the moment people enter the room until CO_2 meter can detect any changes. But the presence of the people in a room can change the number of detected MAC addresses. Benefitting from WIFI sensors to control ventilation system could be more effective as it relies on number of occupants, not the surrounding conditions. Now with the collected data, we can check the opportunities of energy saving by comparing the energy consumption of predefined fixed ventilation and WIFI based one.

We used hourly temperature data provided by Canada governmental website to find the hourly temperature of Toronto city at the time of our experiment (Canada.ca 2017). Table 2 shows an example of the aforementioned hourly temperature data for Toronto.

Table 2- Hourly data report, Toronto, 2017.

	A	B	C	D	E	F	G
17	Date/Time	Year	Month	Day	Time	Data Qual	Temp (°C)
18	2/1/2017 0:00	2017	2	1	0:00		-2.8
19	2/1/2017 1:00	2017	2	1	1:00		-2.9
20	2/1/2017 2:00	2017	2	1	2:00		-3.1
21	2/1/2017 3:00	2017	2	1	3:00		-2.6
22	2/1/2017 4:00	2017	2	1	4:00		-1.3
23	2/1/2017 5:00	2017	2	1	5:00		-1.1
24	2/1/2017 6:00	2017	2	1	6:00		-0.8
25	2/1/2017 7:00	2017	2	1	7:00		-0.2
26	2/1/2017 8:00	2017	2	1	8:00		-0.1
27	2/1/2017 9:00	2017	2	1	9:00		0.1
28	2/1/2017 10:00	2017	2	1	10:00		1.3
29	2/1/2017 11:00	2017	2	1	11:00		1.6
30	2/1/2017 12:00	2017	2	1	12:00		2.1
31	2/1/2017 13:00	2017	2	1	13:00		2.1
32	2/1/2017 14:00	2017	2	1	14:00		3.1
33	2/1/2017 15:00	2017	2	1	15:00		2.5
34	2/1/2017 16:00	2017	2	1	16:00		2.2
35	2/1/2017 17:00	2017	2	1	17:00		1.7

Figure 14 represents the variation of temperature based on hourly data report of Canada website.

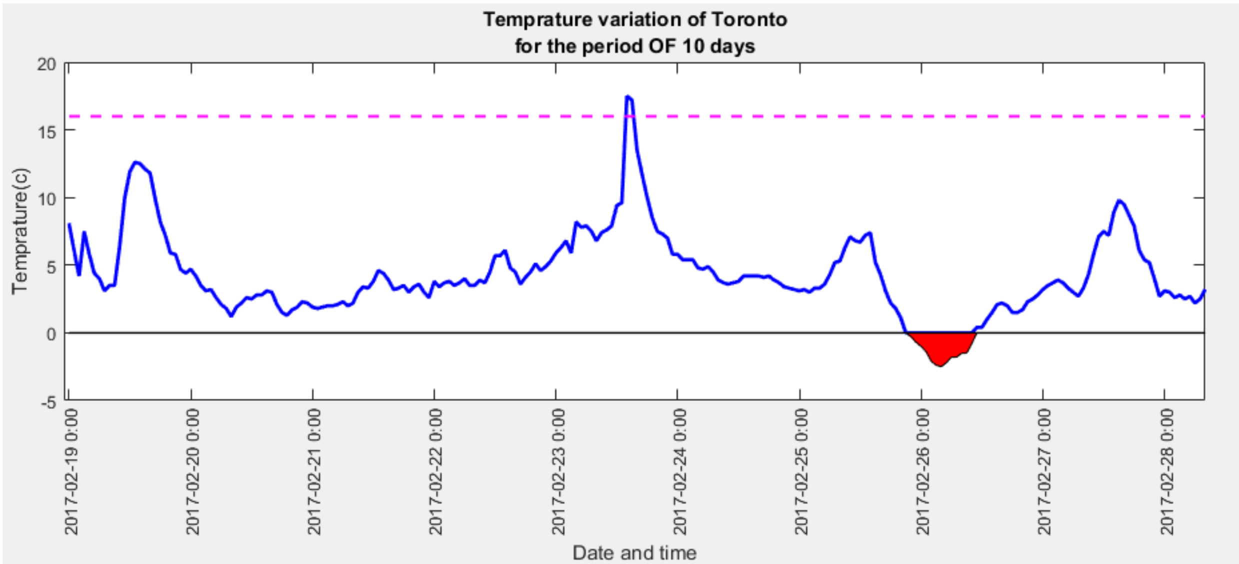

Figure 14. Temperature variation, February, 2017.

As illustrated the February was not as cold as previous years. There was just one day with below 0 temperature. Pink dashed line at 16 ºC shows the ventilation set point temperature. The ventilation system has to consume energy to change the outdoor temperature to this desired temperature. The set point has been selected based on ASHRAE 2013 standards.

According to ASHRAE 62.1 reference, there could be two ways to ventilate the desired area. It can be either based on the total area of the space or the number of occupants [8].

The below table shows the minimum ventilation rate based on the occupancy category. For example, in office spaces, outdoor air rate is 5 Cubic Feet per Minute (CFM) per person or 2.5 Liter per second per person. If it is based on the area of the building it requires 0.06 cfm/ft^2 or 0.3 L/s.m^2. The default values on the right of the table should be used when the actual occupancy is unknown.

Table 3. Minimum ventilation rate for different occupancy categories

TABLE 6.2.2.1 Minimum Ventilation Rates in Breathing Zone *(Continued)*
(This table is not valid in isolation; it must be used in conjunction with the accompanying notes.)

Occupancy Category	People Outdoor Air Rate R_p		Area Outdoor Air Rate R_a		Notes	Default Values: Occupant Density (see Note 4)	Default Values: Combined Outdoor Air Rate (see Note 5)		Air Class
	cfm/ person	L/s· person	cfm/ft^2	L/s·m^2		#/1000 ft^2 or #/100 m^2	cfm/ person	L/s·person	
Coffee stations	5	2.5	0.06	0.3		20	8	4	1
Conference/meeting	5	2.5	0.06	0.3		50	6	3.1	1
Corridors	—	—	0.06	0.3		—			1
Occupiable storage rooms for liquids or gels	5	2.5	0.12	0.6	B	2	65	32.5	2
Hotels, Motels, Resorts, Dormitories									
Bedroom/living room	5	2.5	0.06	0.3		10	11	5.5	1
Barracks sleeping areas	5	2.5	0.06	0.3		20	8	4.0	1
Laundry rooms, central	5	2.5	0.12	0.6		10	17	8.5	2
Laundry rooms within dwelling units	5	2.5	0.12	0.6		10	17	8.5	1
Lobbies/prefunction	7.5	3.8	0.06	0.3		30	10	4.8	1
Multipurpose assembly	5	2.5	0.06	0.3		120	6	2.8	1
Office Buildings									
Breakrooms	5	2.5	0.12	0.6		50	7	3.5	1
Main entry lobbies	5	2.5	0.06	0.3		10	11	5.5	1
Occupiable storage rooms for dry materials	5	2.5	0.06	0.3		2	35	17.5	1
Office space	5	2.5	0.06	0.3		5	17	8.5	1
Reception areas	5	2.5	0.06	0.3		30	7	3.5	1

According to the floor plan of YNG 7th floor, our area of interest is equal to 387 m^2 (Figure 10). There are 2 approaches to design the ventilation rate for offices with different sections. I.e. lobbies, receptionist, corridors. The first approach is to assume the entire zone as office spaces. The other approach is to calculate each part separately. We decided to go with the first approach as subzones of an office area has the same minimum ventilation rate ASHRAE 2013.

The minimum ventilation rate in breathing zone is calculated using below formulas [57]

$$V_{bz} = R_p \times P_z \tag{7}$$

or

$$V_{bz} = R_a \times A_z \tag{8}$$

Where,

V_{bz}= Minimum ventilation rate (breathing zone outdoor airflow)

A_z = floor area of the ventilation zone (ft^2, m^2)

P_z = number of occupants in the ventilation zone

R_p = outdoor airflow rate required per person(CFM per person)

R_a = outdoor airflow rate required per unit area(CFM per f^2)

Floor area can be extracted from the floor plan and the outdoor airflow rate can be obtained from ASHRAE standard [8]. The unknown parameter here is P_z , which refers to the number of occupants.

The area of our office space is 387 m^2 or 4165.5 ft^2. Based on ASHRAE 62.1 standards, the required area outdoor air rate (R_a) for this area is equal to 0.06 cfm/ft^2. Therefore, the V_{bz} for our office area is equal to 249.93 cfm (4165.5 ft^2 * 0.06 cfm/ft^2).

Since we want to calculate the minimum ventilation rate based on actual occupancy, we used the YNG MAC address data as an indication of the number of occupants. For instance, the required outdoor air rate is equal to 5 cfm per person. If there are 20 people in ventilation zone, the V_{bz} is equal to 100 cfm (5 cfm/person * 20). Which is so low compared to the area – based ventilation rate.

The following figure shows cfm rate variation for 1day based on the number of occupants. The maximum cfm needed is 125. As it can be seen, required air flow fluctuated a lot during the day as it is related to the occupancy. It was in the lowest before 7 A.M, then it started to rise. The maximum needed was before noon. During lunch, it decreased a little as occupants may leave the office for lunch. After 4 P.M the required air flow lowered and reached its minimum value due to the ending of the working day (Figure 15).

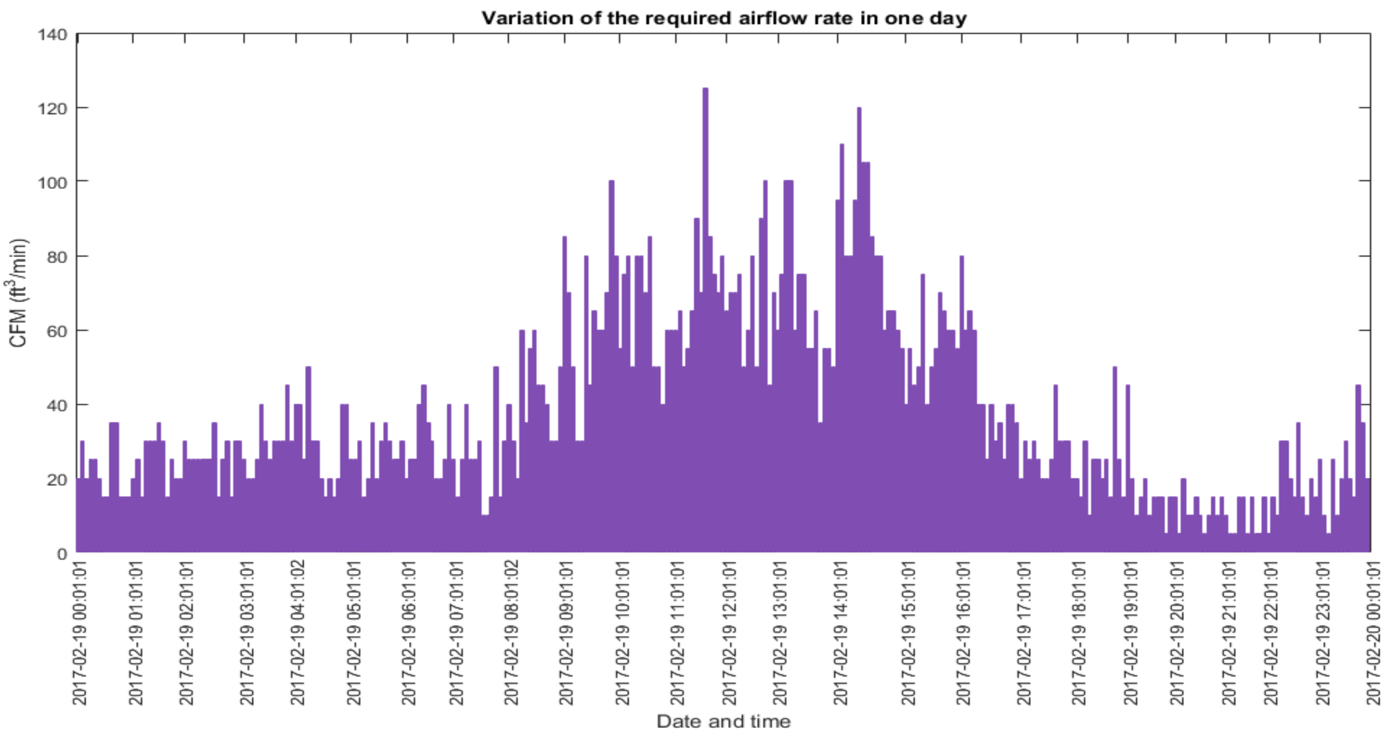

Figure 15. Airflow rate variation in one day.

Calculating the required ventilation airflow base on the office area would lead to the much higher value of 249.93 cfm. This shows how much energy could have been saved if the ventilation was operating based on the population of that area.

3.1.3 *Ventilation energy consumption reduction*

To calculate the amount of energy reduction, the energy consumption of conventional-designed ventilation system has been compared to the occupancy based one. The conventional ventilation system works from 7 A.M until 7 P.M. For better comparison, we also consider this day time period in our calculations for the new design.

The energy that is needed in order to change the outdoor air temperature to the desired set point is being calculated using equation (5).

$$Q = \mathrm{mc}\Delta\mathrm{T} \tag{9}$$

$$Q_V = V_{bz} * \mathrm{m} * \mathrm{c} * (T_{sp} - T_{sa})/e \tag{10}$$

Where,

Q = heat energy lost or gained (J)

Q_v= the energy consumption per unit time to condition the airflow

V_{bz}= minimum ventilation rate (outdoor airflow (cfm))

m = mass of ft^3 of air (0.034 Kg)

C = specific heat capacity of air (1.004 kJ/kg.C)

cfm= volume of airflow ft^3/min

ΔT= change in temperature

T_{sp} = set point temperature

T_{sa}= supply air temperature

e = system efficiency (here we assume it is equal to 80%)

Consequently, if the number of occupants decreases the airflow volume will be less than the fixed pre-scheduled design. Which results in lower energy usage.

Our data has been gathered at 5 min intervals. Since the energy formula and temperature data is hourly based, we considered the maximum occupancy at each hour.

The following figure shows the difference between the original design and occupant-based ventilation energy consumption. As it can be observed the difference is significant. The green plot shows the original design, although the ventilation zone has not changed, the hourly energy consumption has changed due to temperature fluctuation. The hourly average of energy consumption for the area-based design is almost 1.95 KWh.

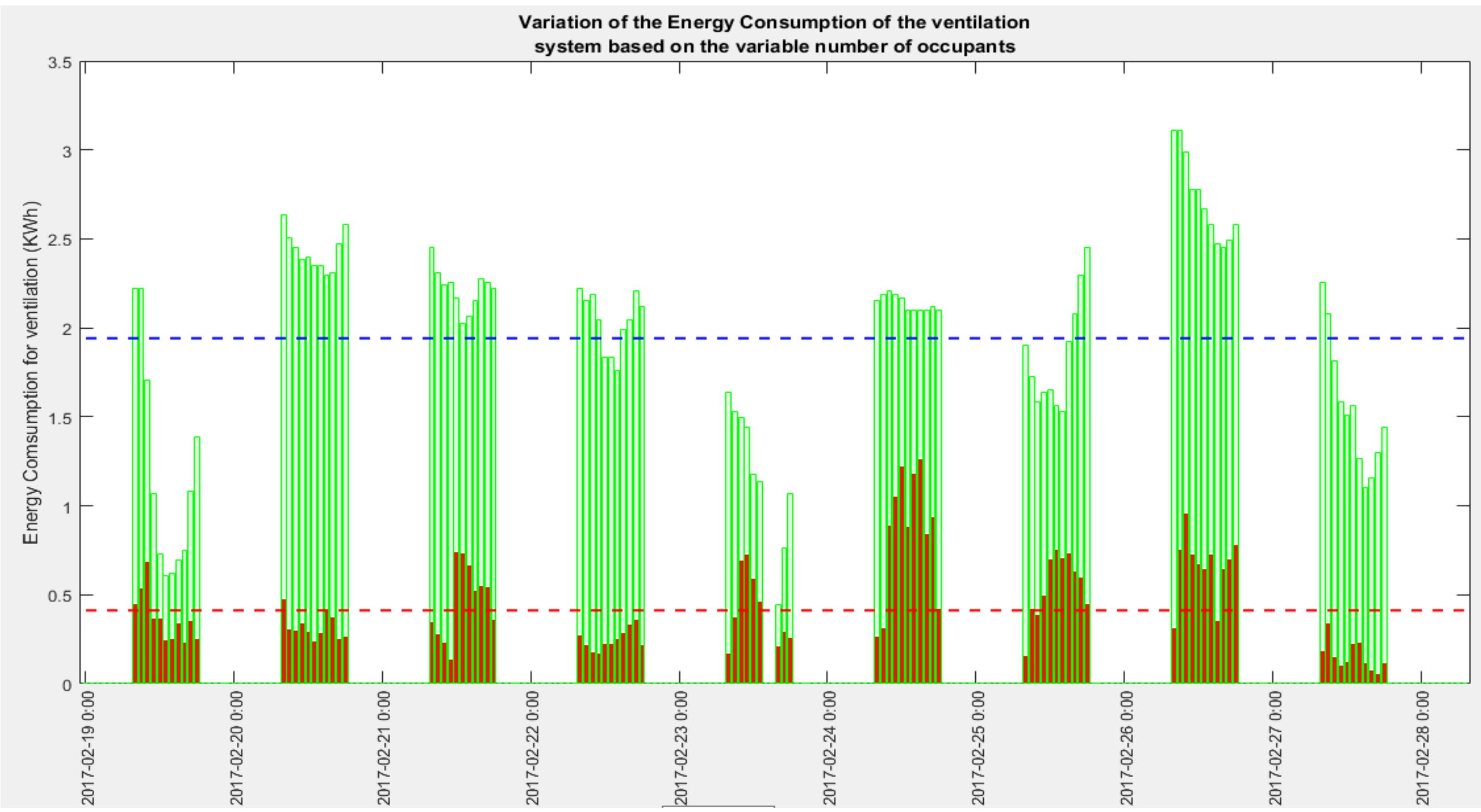

Figure 16. Comparison of ventilation energy consumption o in original design and the occupant-based design

On the other hand, in the occupant-based ventilation design, the average value decreases to 0.41KWh. There is one missing bar between February 24th and February 25th. That's the result of a temperature higher than the set point, which is visible in Figure 16. Therefore, there was no need for ventilation to increase the temperature to the set point.

Total energy consumption for the original design and occupant-based for the experiment period is equal to 209.73 (kWh) and 44.53 (kWh) respectively. That is equal to 78% reduction in energy consumption.

In the case that the ventilation works all day long (24 hours), the energy consumption would differ according to the following figure.

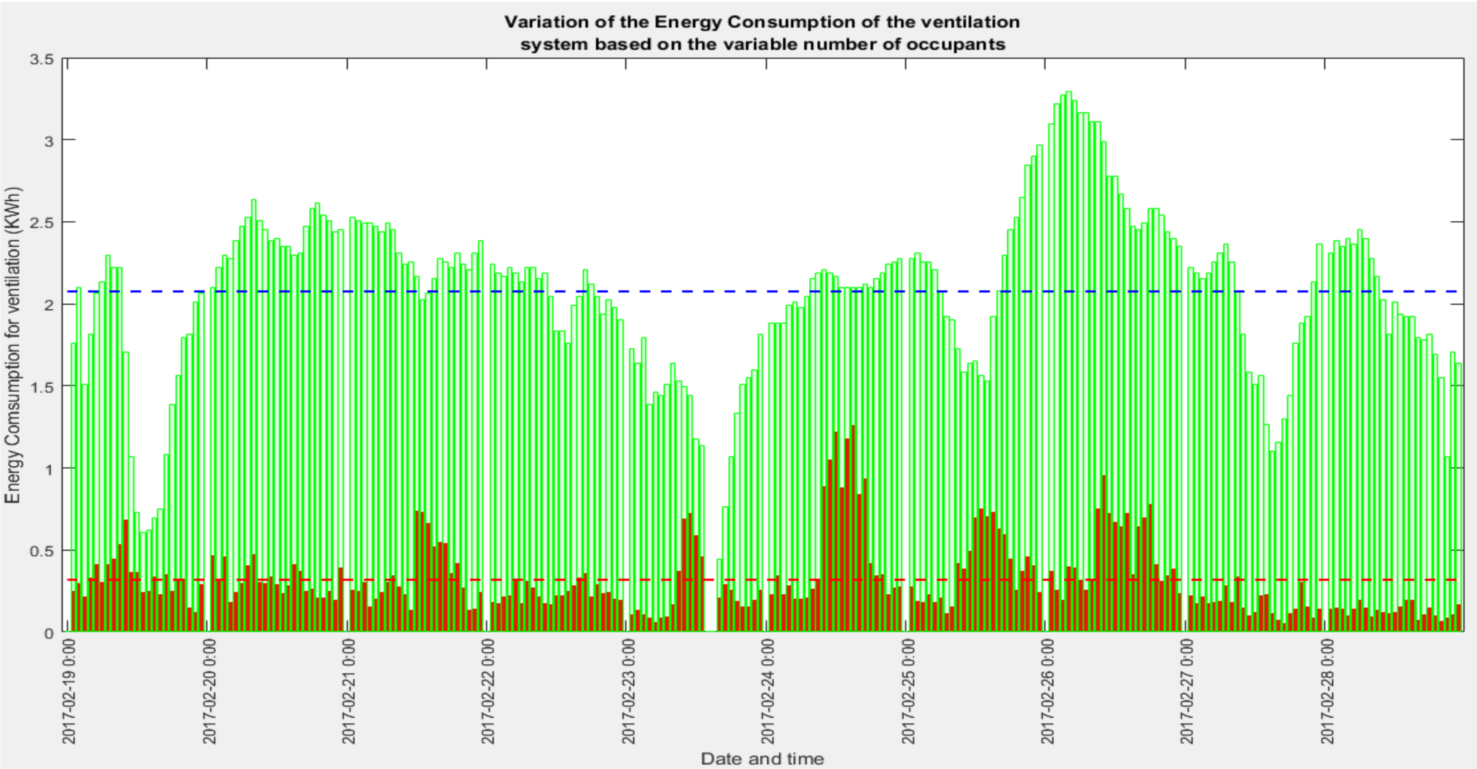

Figure 17. Energy consumption comparison between original design and occupant-based design (24-hour ventilation)

In this scenario, the difference in energy consumption would be more visible. The average is equal to 2.08 (kWh) and 0.32 (kWh) in original design and occupant driven design respectively (Figure 17).

The total energy consumption for the experimental period is 473.18kWh in the original design and 72.61kWh in the new design. This means 85% reduction in energy consumption of the system which could be even more sensible if we consider the ventilation system energy consumption over a long term period.

This year compared to previous years Toronto had a mild winter in February in which there were just a few days with below zero temperature. Therefore, we decided to compare our result with previous years' temperature as well to better test the new method on a typical Canadian winter condition.

Since we collected the data for 21 days in our experimental period, to reduce irregularities within each week, we decided to average over all three weeks and consider it as a nominee of a typical week. Then we extended this week to the whole month of February as it is reported to be the coldest month of a year in Canada.

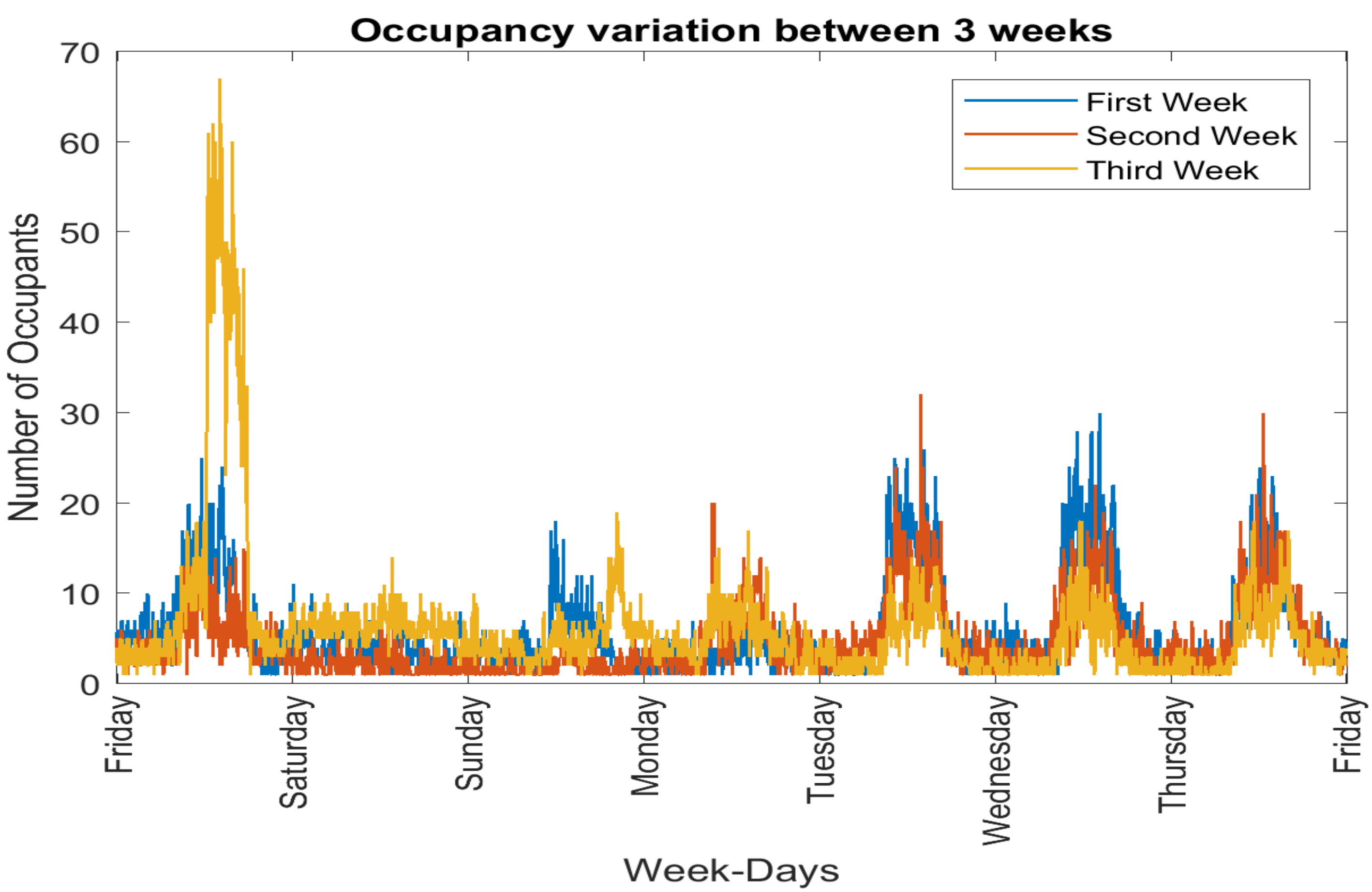

Figure 18. Occupancy variation between 3 weeks.

Figure 18 represents the occupancy pattern in 3 weeks. Different weeks are denoted by different colors. On Friday of the third week we observed an unordinary number of occupants. This clearly shows a predefined fixed ventilation system cannot meet the requirements in unusual situations. Other days follow almost a similar pattern.

Figure 19 illustrates the average occupancy over 3 weeks. Which is below 20 from Tuesday to Saturday and below 13 from Saturday to Monday. We compared the energy consumption of the ventilation system for February between the years of 2014 until 2017 to better observe the energy consumption reduction. The weather data had been retrieved from Canada website [58].

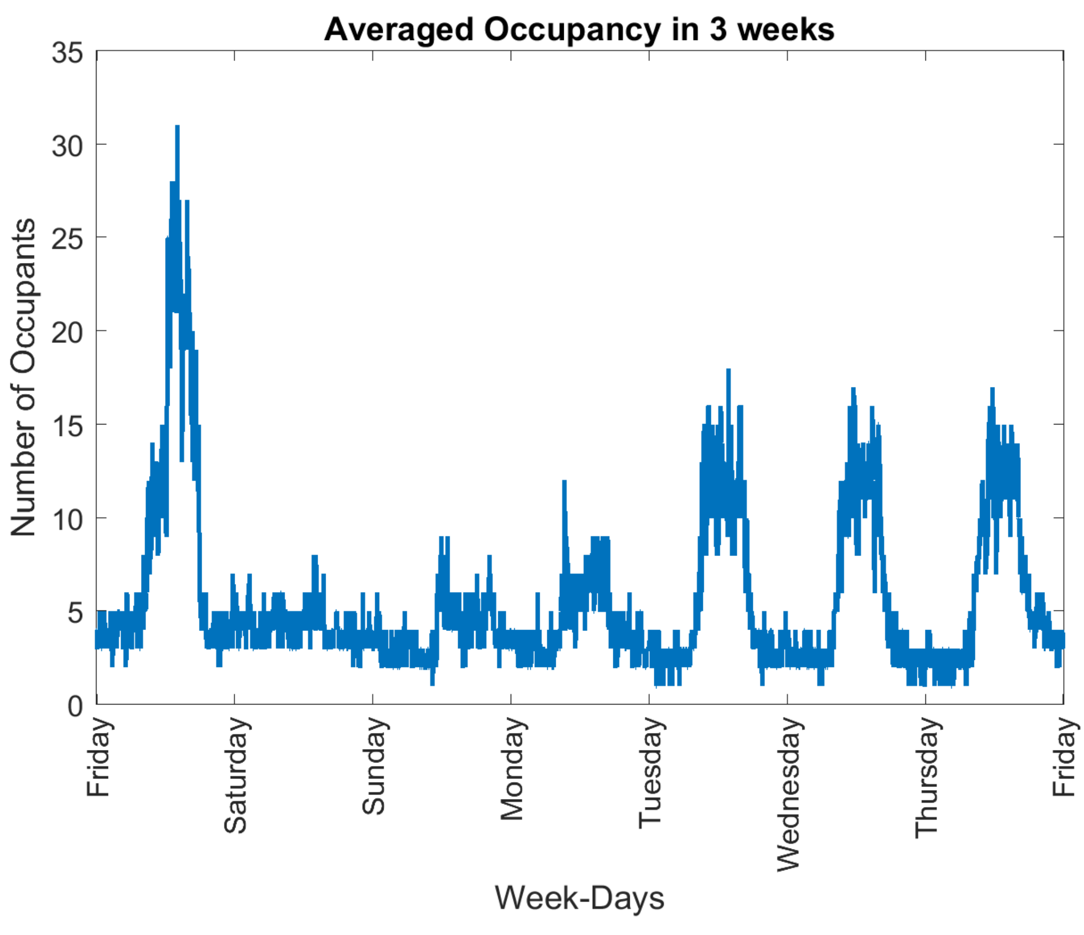

Figure 19. Average occupancy in 3 weeks.

Figure 20 represents the energy consumption of February 2017. We assumed the ventilation system worked from 7:00 A.M until 7 P.M. The area based and occupant based ventilation energy consumption has shown by green and red respectively. The difference is so clear especially on cold days. As shown, the energy consumption has reached near to 5 kWh on February 9th as it was a quite a cold day (-10ºC). But in total this month was not so cold compared to previous years.

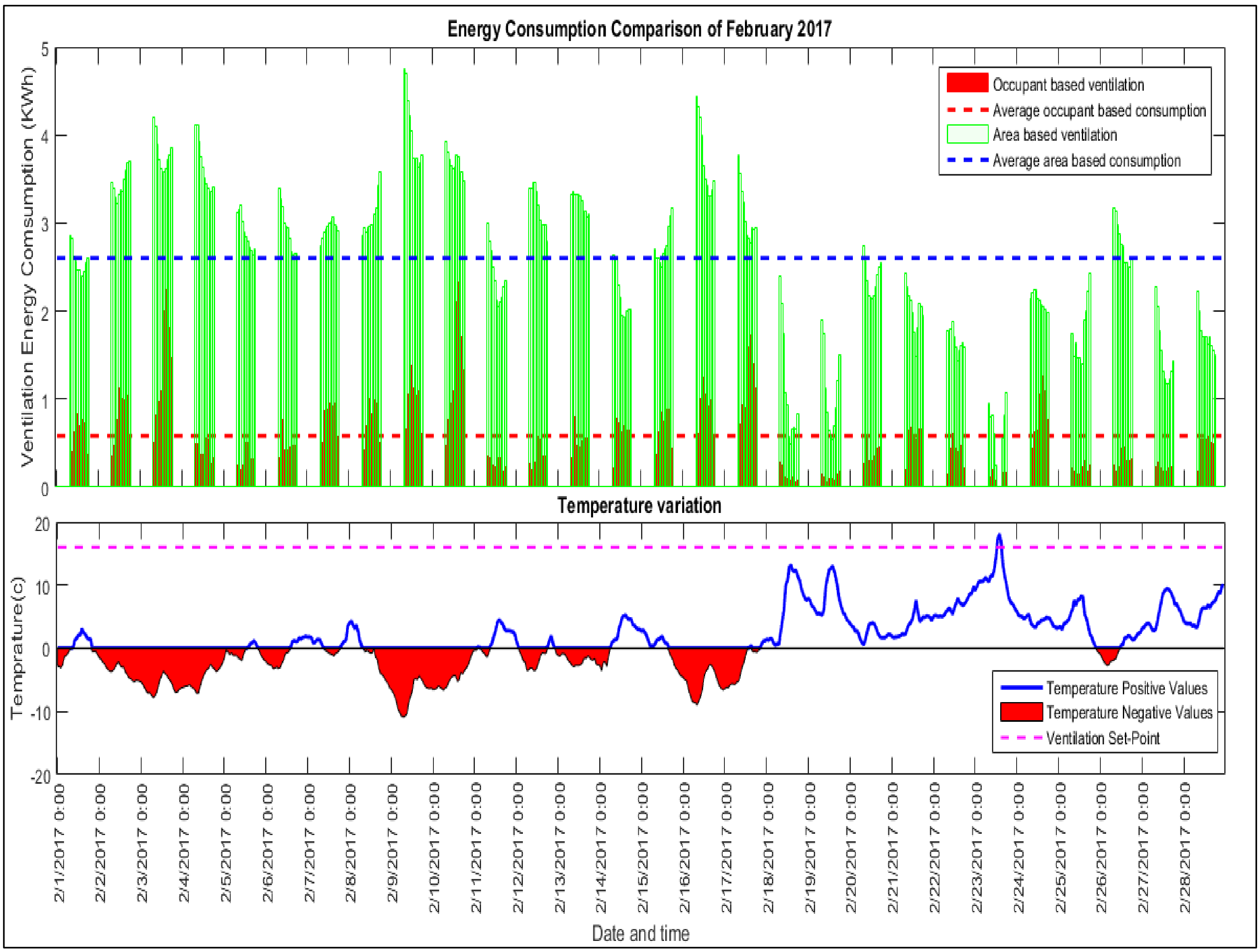

Figure 20. Energy consumption comparison of ventilation system (12Hr), February 2017.

Opposed to 2017, Canadians experienced a very cold February in 2015. The temperature was never above 0 ºC. This makes ventilation system to spend too much energy to bring the temperature to the specified set point (16 ºC). That is why energy consumption has increased proportionally to the temperature. The average consumption is around 5 kWh which is almost twice as that on 2017. The average usage of area-based ventilation is 5 times more than occupant one (Figure 21).

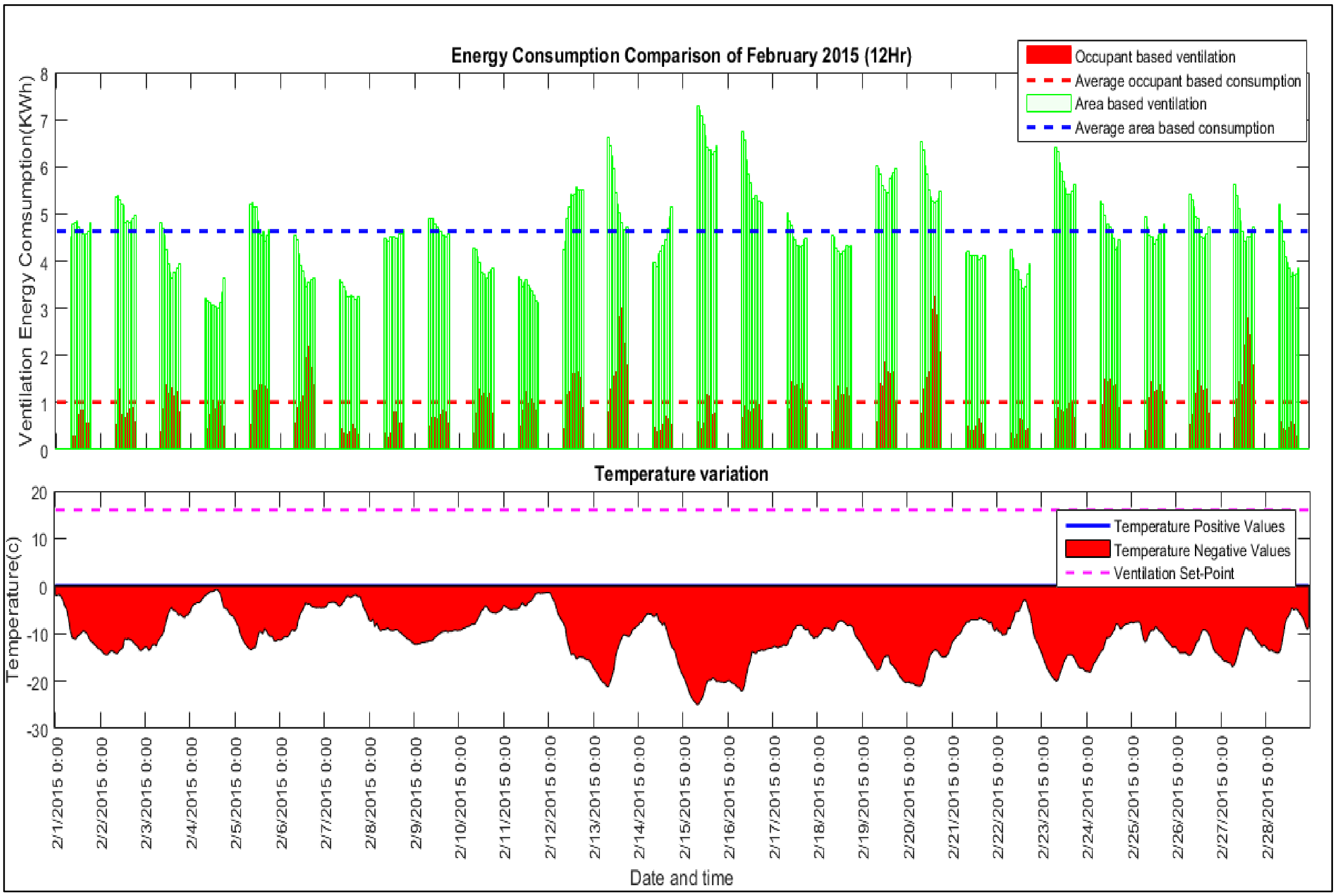

Figure 21. Energy consumption comparison of ventilation system (12hr), February 2015.

Related figures for the year 2014 and 2016 along with 24Hr designs are located in the appendix (Figure 34 - Figure 37).

The comparison of energy consumption of February between 2014-2017 has been gathered in the following figure. We assumed ventilation had been working for 12 hours starting 7 A.M until 7 P.M (Figure 22).

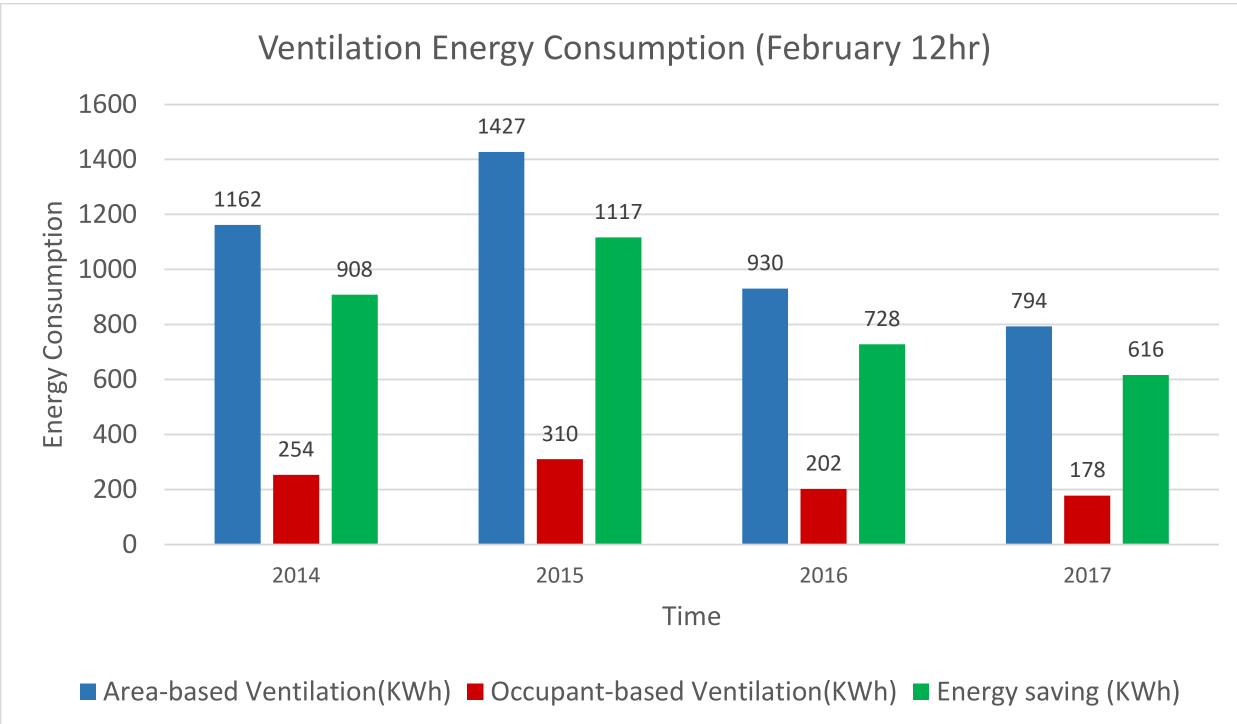

Figure 22. Ventilation energy consumption comparison (2014-2017) 12hr design.

As stated, energy saving is more evident on cold days. 2015 which was the coldest among recent years had the opportunity of 1.117 MWh energy saving with the new ventilation system and only for the month of February.

According to ASHRAE guideline, for the cases that occupants may stay over night or after working hour, ventilation system should also operate. This is something that is being neglected in most of the conventional ventilation designs. In the 24-hour operation mode, the occupancy-based ventilation saves much compare to 12Hr mode setting. In Figure 23, we observe huge savings, especially in 2015.

Comparing Figure 22 and Figure 23 the new ventilation design consumption has not been changed significantly compared to 12-hour design. For example, in 2017 it consumed 178 and 265 kWh in 12hr and 24hr respectively. But the old design used almost 1000 kWh more in 24hr design (Figure 23).

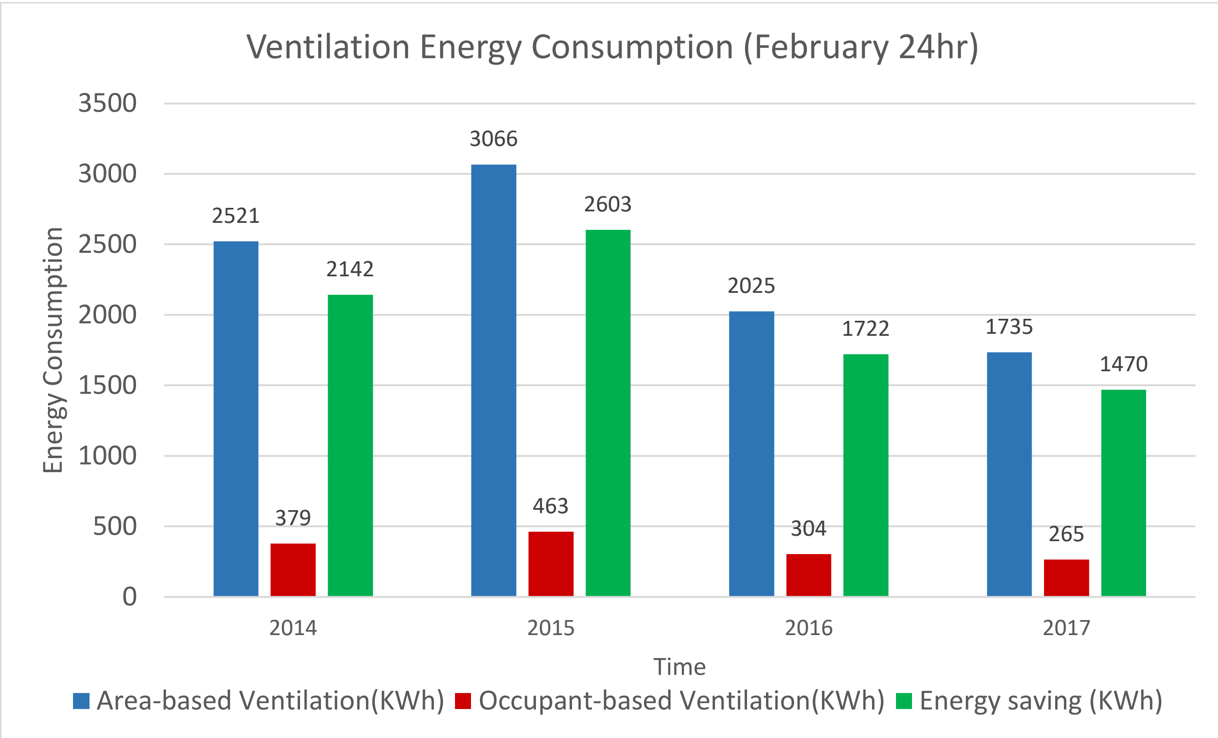

Figure 23.Ventilation energy consumption comparison (2014-2017) 24hr design.

3.2 Indoor localization in ventilation system optimization

There are different methods to determine the indoor position of the client such as wearable sensors but these methods need extra cost and installation or configuration. Indoor positioning which is based on RSSI compared to the other methods does not need extra configuration and cost as it can be deployed on the existing WiFi infrastructure. Since the RSSI nature is unpredictable and it varies over the time, fingerprinting method has been used in this thesis to overcome the RSSI challenges and locate the occupants accordingly. Our first intention was to make use of already existed access points of Ryerson University to find the position of occupants in the building without using extra sensors and software. Because this part required too much of CCS staff time and direct access to the building access points was not possible due to the security matters, therefore we decided to use INSSIDER software to monitor and record RSSI data. The zone of interest in this experiment is office 700 in YNG building located at Yong Street.

3.2.1 *Challenges*

Due to the unstable nature of the RSSI, there are some challenges for the RSSI-based indoor positioning. The RSSI will change over time due to the changes in the building structure, doors, partitions, etc. other sources such as Bluetooth and microwave can interfere with the RSSI reading. Water is the main component of the human body that can absorb the WiFi signals. Therefore, human presence will affect the received signal. Moreover, the direction of the device will influence the received signal [59]. This variation in RSSI can be addressed by using a huge fingerprint database that has been gathered during different time stamps. This database must be updated if some major changes happen at the test location.

3.2.2 *Fingerprinting*

A fingerprint indoor positioning system consisted of 1) three or more access points 2) a device that can collect RSSI from the APs; 3) a server to store the RSSI data collected by the mobile device; 4) an algorithm that can use the database to estimate the location.

3.2.3 *INSSIDER*

For training phase, INSSIDER which is an open source WiFi scanner has been used. This software can be operated both on desktop and mobile phone. In our case, we used the windows version. INSSIDER scans all the existing access points every second and log the data at the same time in .gpx format (Figure 24)

The scan consisted of SSID, Mac address and RSSI of each access point.


```
<gpx>
  <wpt lat="" lon="">
    <ele>0</ele>
    <time>1-01-01T00:00:0.0Z</time>
    <geoidheight>0</geoidheight>
    <name>RU-Secure [18:64:72:DF:F0:03]</name>
    <cmt>0</cmt>
    <desc>RU-Secure
[18:64:72:DF:F0:03]
RSSI: -67 dB
Quality: 94%
Channel 1
Speed (kph): 0
1-01-01T00:00:0.0Z</desc>
    <fix></fix>
    <sat>0</sat>
    <hdop>0</hdop>
    <vdop>0</vdop>
    <pdop>0</pdop>
    <extensions>
      <MAC>18:64:72:DF:F0:03</MAC>
      <SSID>RU-Secure</SSID>
      <RSSI>-67</RSSI>
```


Figure 24. The raw data of INSSIDER software

3.2.4 *Phase one, training mode*

As occupants in an office are mostly seated at their desks, their position behind the desk was taken as the reference points. The INSSIDER equipped laptop was positioned at each reference point and the scan ran for 20 minutes at each point.

During the offline phase, the area of interest has been segmented into 8 sections (8 desks). The laptop gathers the RSSI values $\varphi_{i,j}^{\theta}[r]$ from each AP at each reference point. Since the orientation of the device will affect the RSSI value the RSSI was gathered parallel to the 4 main directions of human movement in the room. The radio map is created as below:

$$\Psi^{\theta} = \begin{pmatrix} \varphi_{1,1}^{\theta}[r] & \varphi_{1,2}^{\theta}[r] & \dots & \varphi_{1,n}^{\theta}[r] \\ \varphi_{2,1}^{\theta}[r] & \varphi_{1,2}^{\theta}[r] & \dots & \varphi_{2,n}^{\theta}[r] \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{b,2}^{\theta}[r] & \varphi_{b,2}^{\theta}[r] & \dots & \varphi_{b,n}^{\theta}[r] \end{pmatrix} \tag{11}$$

Which represents the RSSI value measured from i^{th} access point at the j^{th} reference point with the orientation of θ. b is the number of access points and n is the number of reference points. r = 1,.., N and it denotes the number of samples.

Moreover, as RSSI will change over time. For better estimation, the data collected at different time stamps at 3 different days: 2 days for training and 1 day of testing. The data then transferred to MATLAB, restructured and unnecessary information was deleted. As we wanted to examine our fingerprinting method in a heterogeneous environment, 5 access points of Ryerson University and 4 other existed access points were selected for the experiment.

Table 4 Excel format of the inssider data after being organized

RU_SECURE	RU_SECURE	RU_SECURE	RU_SECURE	RU_SECURE	AP01	AP02	AP03	AP04
[18:64:72:DF:BE:23]	[18:64:72:DF:CA:63]	[18:64:72:DF:F0:03]	[6C:F3:7F:EF:F1:63]	[94:B4:0F:05:1E:E1]	[C4:12:F5:7C:83:46]	[C4:12:F5:7C:66:80]	[C4:12:F5:7C:70:C8]	[C4:12:F5:7C:83:2A]
-69	-82	-58	-61	-47	-46	-43	-49	-61
-69	-82	-57	-61	-47	-45	-52	-47	-58
-69	-82	-57	-61	-47	-45	-52	-50	-58
-69	-82	-57	-61	-47	-45	-45	-56	-63
-69	-82	-57	-61	-47	-45	-56	-49	
-69	-82	-57	-61	-47	-45	-56	-49	
-69	-82	-58	-63	-45	-48	-46	-64	
-69	-82	-59	-63	-45	-52	-44	-50	-65
-69	-82	-59	-63	-45	-95	-45	-50	-95

As it can be observed, although the position of the laptop has not changed the RSSI value fluctuates. These changes have been smoothly varied in the case of Ryerson access points as these AP are more powerful than the average access points. For the rest of the access points, RSSI has changed significantly.

Moreover, during the scan for some access points, there were no readings for a few seconds. Therefore, it was left empty in the RSSI matrix TABLE 4. In the upcoming sections, these empty cells were treated in two different processes and the results were discussed.

3.2.5 *Phase two, online mode*

During this phase, the device will go around in the desired area and record RSSI from each node. The online RSSI will be compared to the database and the after finding a match it will be marked as the location of the occupant. Although the data must be the same as the database, due to noise and other interferences the values will be different. We used Artificial Neural network to find the indoor location of the wifi device.

Pre-processing techniques such as normalization of data was not applied in our system as the nature of RSSI is non-linear. Changes in the received signal were due to movement of other occupants and the direction of the device. The goal is to build a fast none complex pre-processing system.

3.2.6 *Artificial Neural network*

ANN is a program that will use the input and output data to train itself. The network will be trained more efficiently as the input data for training increases, though the inputs must be varied for different situations. If the train set is too long and repeated but lacks the generalization, the result could be local optima and not universal [60].

The performance of the ANN depends greatly on the number of layers, number of neurons and the weight of each neuron. There is no general rule to determine the best combination of these factors. In fact for each experiment, number of layers and neurons may vary significantly based on the problem. Each time you train the NN, different initial weight and a different portion of data will be used. Therefore, the output may change at each attempt.

Mehmood et al. used ANN to determine the indoor position. They used RSSI data and then calculate mean, max of captured RSSI at each reference point and used these values as the input of their network [59]. They have not checked the degradation after a few days of database creation. They reported 30% error within 1m and 60% of distance error within 1-2 m.

3.2.6.1 *Number of Layers*

All neural network architectures have one input layer and one output layer. The number of neurons in each layer depends on the training dataset. In this thesis, we used 8 access points which result in 8 input variables as columns and the rows of the input data are equal to the number of samples. We changed the number of samples and compared the result. Since our data is going to be classified in 8 different reference points the number of output is 8. (Figure 25)

The layers that rest between the input layer and output layer are called hidden layers. In this thesis, we used just one hidden layer with a number of neurons.Although many researches had been conducted to set a rule to estimate the best number of neurons in hidden layers, there is no rule of thumb to calculate the optimal number of them. Some researchers have mentioned the following rules to achieve the best number of hidden neurons [61]:

- The number of neurons in hidden layer is equal to:

(Number of inputs + number of outputs)$\times \frac{2}{3}$

- Input neurons $<$ hidden neurons $<$ Output neurons
- The number of hidden neurons $<$ 2 (number of input neurons)

If the number of neurons is not enough for the data set it will cause under the fitting problem. On the other hand, too much hidden neurons will result in over fitting problem. In both cases, the training dataset would show inaccurate predictions over the test data set.

The optimal number of hidden neurons for this thesis has been determined by error and trial and comparing the result for each change.

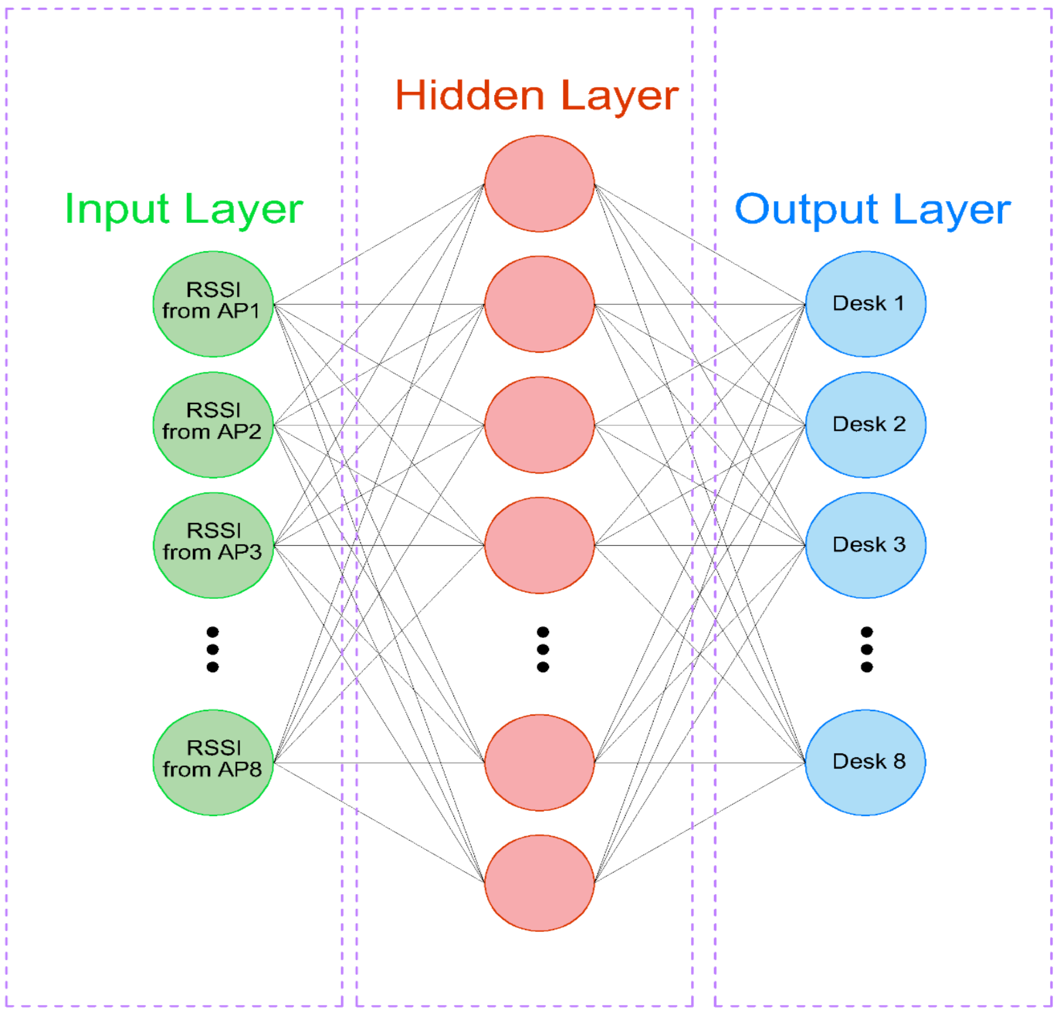

Figure 25- neural network architecture used in this thesis

3.2.6.2 *Training algorithm*

For training the data in the supervised method you have to give the network a set of input and the correct output. There are different ways to train a network. Trainlm (Levenberg-Marquardt) is one of them which is fast. Though, its performance degrades as the data size increases. Trainscg or the Scale Conjugate

Gradient is a better choice for pattern recognition as it performs well with larger data and the speed is fast. Trainbr or Bayesian regularization will reduce the squared errors and weights to determine the best set of weight for the network. Bayesian regularization will generate a network that can generalize efficiently. We examined some of the training methods and compared the result.

All training algorithms have a stop condition. Otherwise, training will never stop. The process will usually stop when the model meets the below restrictions:

- It reaches the maximum number of epochs.
- It overrides the time limit.
- The desired Performance meets the set goal.

If the output of the NN is not satisfactory, there are some solutions that can be tested in order to achieve better accuracy.

- Adding more input data.
- Adding more neurons in the hidden layer.
- Testing other training algorithms.

3.2.6.3 *Overfitting problem*

This is a common problem that happens during testing. Once you run your model, the accuracy on the training set is so high. But when you test the model on a new set of data it will drop drastically. This is due to the fact that the network had trained itself well for the first set of data, but lacks generalization for the new set of input. Bayesian regularization is a method that can avoid overfitting.

Bayesian Regularization: This approach changes the performance function which measures the network's performance according to the mean of squared errors.

$$mse = \frac{1}{N}\sum_{i=0}^{n}(e_i^2) \tag{12}$$

By taking into account the mean sum of squared weights (msw) in addition to mse, it is possible to avoid overfitting.

$$msw = \frac{1}{N}\sum_{i=0}^{n}(w_i^2) \tag{12}$$

This method will minimize the value of the assigned weights and makes the NN more robust. The validation set in Bayesian Regularization is not used so that network can use all the input data to train itself. Trainbr is a function that uses Bayesian regularization to train the network. To sum it up, this method works better than early stoppings as it considers all the input for training. However, it takes longer to converge.

3.2.7 *The role of different factors on the results*

3.2.7.1 *Training data set size*

The first factor we used for comparison is the size of our training dataset. The more data being used for training the better accuracy is. However, this input must be varied and includes all the factors. Otherwise, the ANN may perform well on the training input. However, if we test it for unseen data (the data that had not be used at training stage) the result won't be as accurate as expected. That's because of the fact that network lacks generalization due to insufficient sample input. As the number of samples is our main focus in this section, we set the number of hidden neurons to 10 and our training function to 'tainscg'. These two are the default setting in MATLAB ANN toolbox.

For the first examination, we used 100 samples per direction per desk. So it is going to be 100*4*8= 3200 in total.

Then we also used 50 samples per desk per direction to test the data for unseen samples. Below figure shows the result of training.

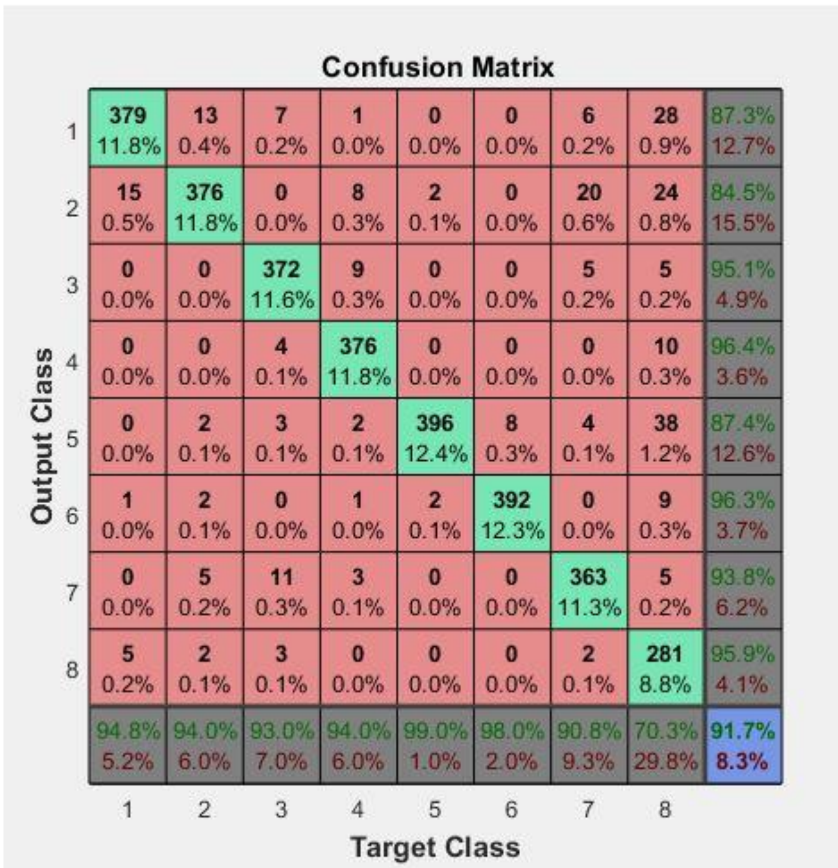

Figure 26. Confusion Matrix for training section with 100 samples

Figure 26 shows 91.7% accuracy for training data set. The Output Class is the network response and Target Class is what we expected to get. Consequently, the green cells show the times that the target was achieved accurately based on the input value. For example in the first row, the input is 1 (desk 1) and 379 of the times the target was selected correctly. Next cell to the right which is red, shows 13 times desk 1 got selected although desk 2 was our goal. The last cell in row 1 shows 12.7% in total desk 1 got selected by mistake.

The first cell in the last row shows 400 times target was supposed to be desk 1, though 379 times network selected that correctly.

Some may think this result is the final result of their network. But the best way to test the network is using some data that had not been used for training as the new input. Figure 27 shows the result of the network for unseen data. As illustrated the result changed by almost 10 percent. This is due to the fact that the network has not been trained well and lacks generalization.

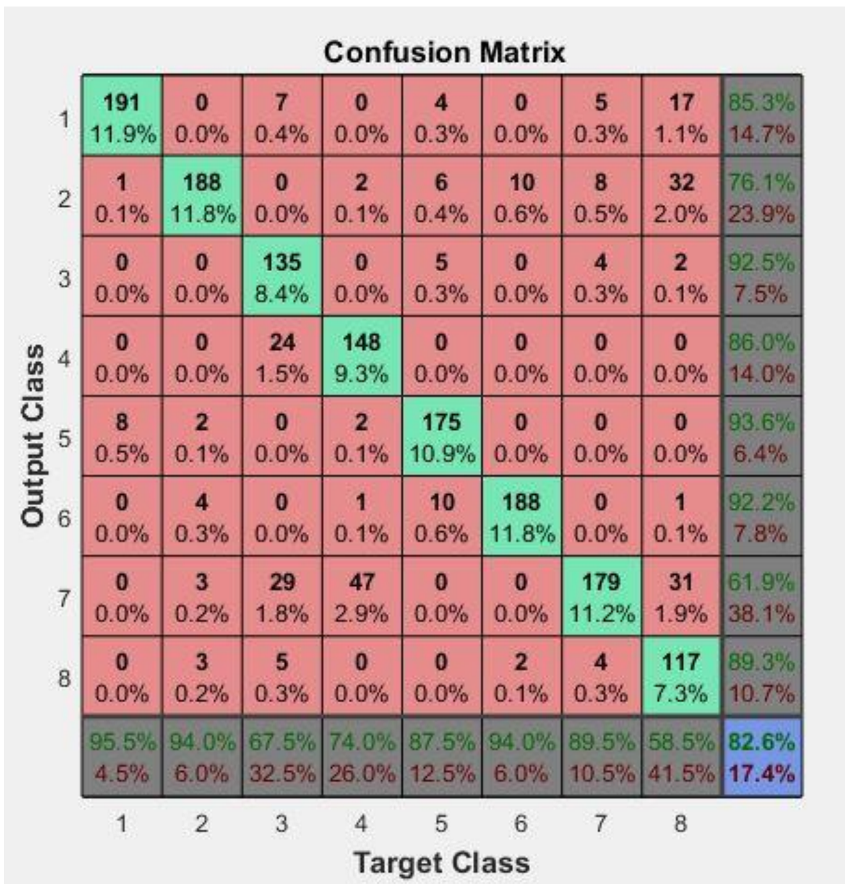

Figure 27. Confusion Matrix with unseen data

For next round of examination, we increased the samples to 300 per direction per desk. Figure 28 shows the confusion plot for training data set. The accuracy changed from 91% (100 samples) to 88% (Figure 28). This means the 100 samples was not enough for training. Therefore, it is advised to consider all possible kinds of inputs to train your network.

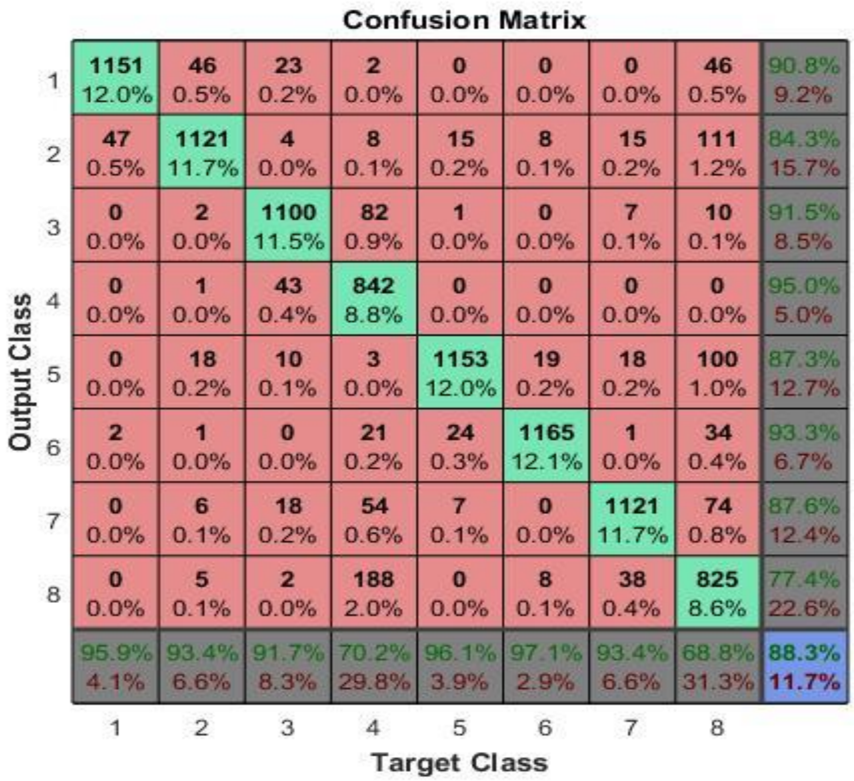

Figure 28. Confusion Matrix with 300 samples

Next, we examined our network with unseen data. Figure 29 presents the result for unseen data. The result in Figure 29 is closer to the training compared to the case with 100 percent. This could be due to the generalization that is more evident in the second network design.

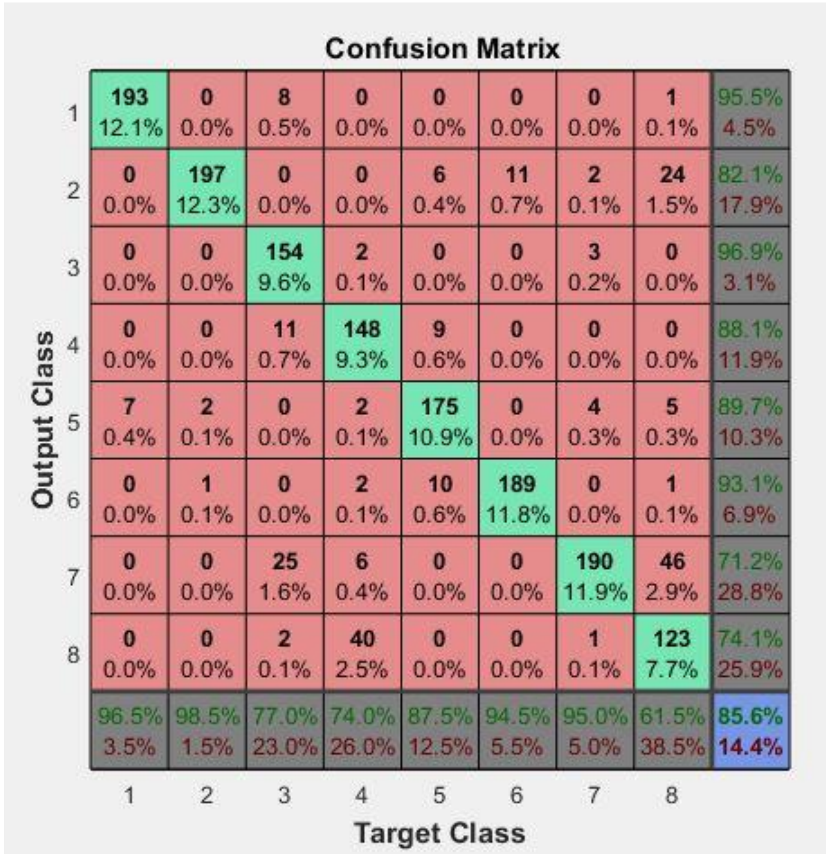

Figure 29. Confusion Matrix for unseen data with 300 samples

We raised the samples up to 500 and 600 but didn't observe so much change in the result. Then we changed the samples to 700 and observed a great improvement (Figure 30).

Figure 30 shows 91.6% accuracy in total. Which is a good improvement. We then checked the network with unseen data to find out how much the result is close to the training. The unseen data shows 89.5% accuracy which is a great number for unseen data. This is due to the fact that the network input is enough and had been generalized well.

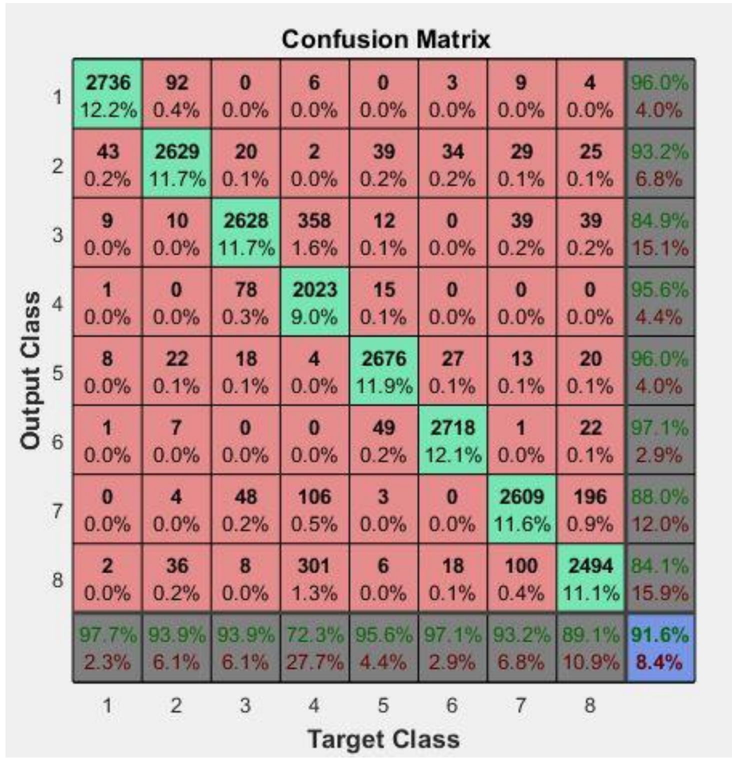

Figure 30Confusion Matrix with 700 samples

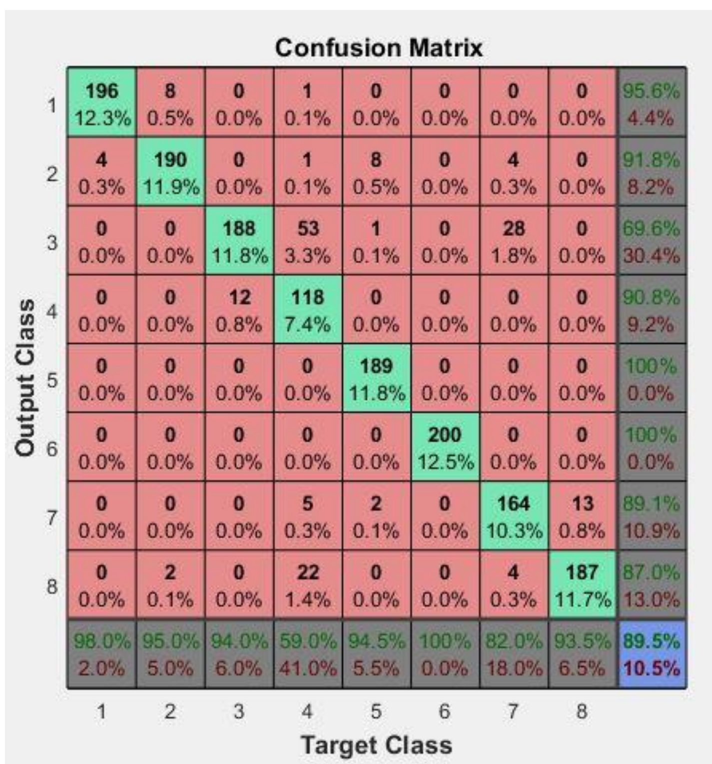

Figure 31. Confusion Matrix with unseen data, 700 samples

3.2.7.2 *Number of Access Points:*

Once we run INSIDER we found 5 access points from Ryerson (Figure 32). These APs (Aruba network) showed a stable pattern of RSSI and the signals were not highly fluctuating over the time, except one of them. The blue access point in figure had a weak signal strength and most of the time there were no readings. -96 indicates no reading and -95 is the weakest RSSI that can be captured with inssider. After an investigation, we found out this AP belongs to the 6th floor which was not in our area of interest. Therefore, we discarded this access point as its data was not helpful.

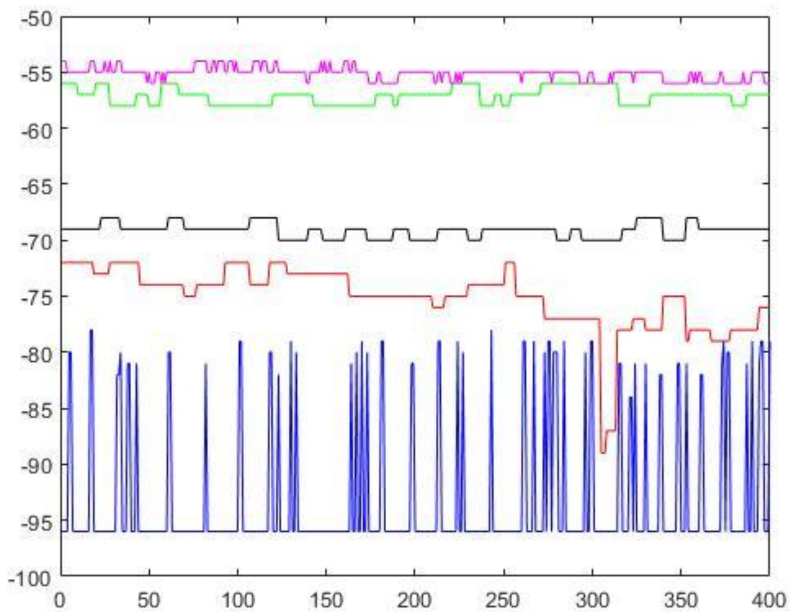

Figure 32. Ryerson Access Points Comparison

As we wanted to use different types of AP in terms of cost and vendor to test our system in a heterogeneous environment. We used 4 other inexpensive, simple AP installed at office 700 (Figure 33).

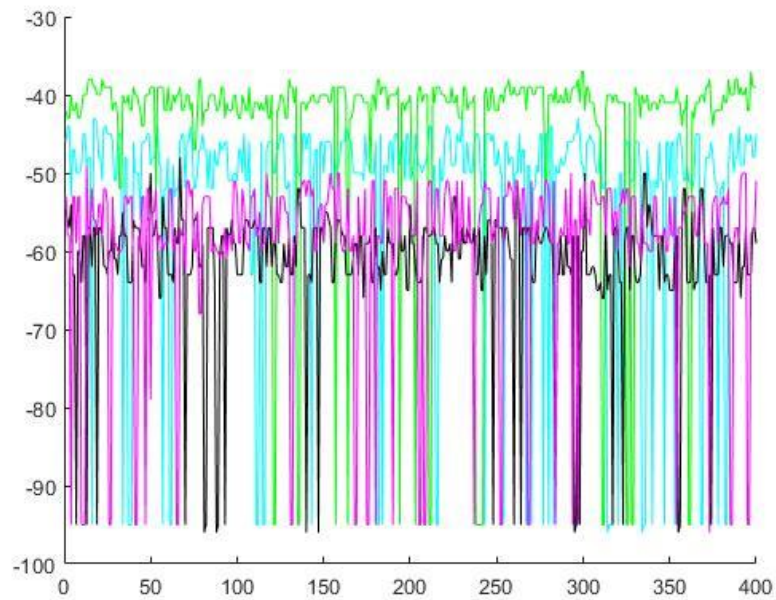

Figure 33. Simple, inexpensive AP comparison

These 4 access points were fluctuating a lot during our experiment. The reason could be related to the amount of noise and occupant movement in the office. These APs are not expensive compared to the ones from Ryerson. Therefore, noise must have a great impact on them. In addition to that, there are a lot of furniture i.e. Desk, chairs, shelves and human movement in the office which can affect the fading and absorption of the signals. Some researchers used places without any furniture to test their system but as we wanted our environment to be a real office, all furniture was remained unchanged. In most of the previous cases the area of study was without occupant movement but during our experiment, the office owners were coming and going as usual.

We altered the number of access points and the combination to see how much it does impact our result. The training function and hidden neurons remained 'trainscg' and 10 respectively. Below table shows our findings.

Table 5. Result of different combination of APs.

number of Ryerson APs (Aruba)	number of Simple APs	Accuracy percentage in training	Accuracy percentage with unseen data
4	0	90.2	83.1
0	4	67.4	60
4	1	90.4	84.2
4	2	90.7	87.3
2	2	70.1	58.8
3	3	80.4	73.6
4	4	91.7	87.8

The first round we used only 4 APs of Ryerson, the result was good for the training part (90.2%) but dropped by 7 percent with unseen data. Next, 4 inexpensive APs got examined. It resulted at 64.4% for training and 60% with unseen data. which is not so good. This is due to the fact that there is huge level of noise and movement in the office. Then we checked different combinations of Ryerson AP and simple APs. The worst one was an equal combination of 4 different type of APs, which resulted in almost 70% and 59% in training and unseen data respectively. Therefore, we decided to keep all 4 APs of Ryerson and gradually add more APs. In training section, not much improvement was observed. But in terms of unseen data results enhanced from 81% to 88%. Thus, we decided to go with 4 Ryerson AP and 4 simple AP. One thing was so clear during our experiment that the quality, vendor and cost of the AP is important in terms of signal strength and noise impact.

3.2.7.3 *Treating zero values*

There were cases where some APs were not detected at specific locations due to nature of the signal. Since all the used APs must have a valid value in the neural network. We treated these empty cells in 2 different ways:

First, we assigned -96 to empty cells which is below the last measurable RSSI of the INSSIDER software (-95). Such as used in [18]. It resulted in 80.3% accuracy at training and 78.1% at unseen data.

The second way was completely deleting the readings in which there was at least one zero value. Although this method would reduce the number of samples, the result showed 90.8% at training and 88% at unseen data testing. This approach had been selected for our system.

3.2.7.4 *Training Method:*

In this section, we checked the effect of some of the training functions and compared the result. The number of training samples and unseen data were considered 500 and 100 respectively. Number of hidden neurons was chosen to be 10. The first method was ‘trainscg’ (scaled Conjugate Gradient). It took 00:04 seconds to converge. 90% and 88% were the correct predictions on training and unseen data. Next, we examined ‘trainlm’ (Levenberg-Marquardt). It converged in 00:10 seconds and resulted in 95% on training data and 89% on unseen data. Using ‘trainbr’ (Bayesian Regularization) takes longer to be converged, 10 minutes in our case but the result was very promising. 96% accuracy in training and 92% in testing. Therefore, trainbr has prevented overfitting problem by a very close result between training and testing.

Chapter 4. Conclusion and Future Work

In this field study, we designed an occupancy-based ventilation system which uses MAC address of occupants' devices as input. This data has been captured through already installed Ryerson University WIFI infrastructure. Contrary to previous occupancy monitoring systems, this design does not need any cost in terms of installation and maintenance. The data pattern was so close to the actual occupancy pattern, which makes the WiFi approach promising. We designed the ventilation rate at YNG building based on the actual occupancy and compared the result with area based ventilation. The ventilation in the original design operates from 7 A.M – 7 P.M and based on the area of the office space. However, using the occupancy-based ventilation design resulted in 78% reduction in 12-hours-operation mode and 85% reduction in energy consumption in 24 hours operation mode.

To further investigate the potentials of WIFI technology to design a zone-level HVAC system we designed a neural network model that uses RSSI value to find the indoor location of the WiFi devices. The data has been gathered in 3 different days at a real office with continues occupant movement and furniture. We test different network architectures and compared the results. The best model achieved 96% accuracy at training data set and 92% at test set. Although data were collected on different days and test data was not being used in training part, the final result did not degrade.

In summary, our results show that occupancy driven ventilation employing WiFi infrastructure can be replaced with the conventional designs. Using this would significantly reduce the ventilation energy consumption.

4.1 Contributions

To investigate the impact of occupants on ventilation energy consumption, two levels of the experiment were considered. In the first one, only the number of occupants in our target zone were monitored. Later it was used to calculate the fresh air demand in the whole zone of interest. In the second level of optimization, we tried to localize the ventilation demand by locating the occupants in subzones. We used indoor localization approach using the artificial neural network in this regard.

Our main goal was to design an occupancy monitoring system by passively using already existed infrastructure to reduce the energy consumption of ventilation system. Although there is a number of researches that leverage WiFi technology to control HVAC system, heating part of this system was the main target of these studies. To the best of our knowledge, this is the first research work that utilizes WiFi infrastructure to control ventilation system airflow rate. A complete demonstration of leveraging existed WiFi infrastructure was given in a real test bed with 78% reduction in ventilation energy consumption.

Commonly, the energy conservation strategies need an initial investment that can be recouped after a few years, but some factors need to be checked to see whether this strategy is economically wise or not. It depends on:

- How long you wish to stay in that building?
- The rise in energy cost
- Does this improvement add to the value of the building?
- The lifetime of the new strategy
- The maintenance cost
- The comfort level after new improvements

In our case, the cost of installation and maintenance of the system was zero as we leveraged already existing WiFi technology.

For the second part of our research work, opposed to previous studies, the idea was to mimic a real office condition with all the furniture intact and continues occupancy movement. We were able to enhance and stabilize the room-level accuracy with 96% in training and 92% in the test set. The data was gathered in different days and different times of the day with multiple occupants present in the room. Most of the previous researches did not investigate the accuracy of their system after the test day [37]. Contrary to the previous studies, the accuracy has not degraded significantly in time [62].

There were no complex post processing or past processing of data required as the proposed neural network model is capable to detect the position with high accuracy. The proposed method is able to be trained efficiently with more data and for bigger areas as Neural Network gets better training with more data.

4.2 Future work

The future of this research area is tremendous and can be further studied. This can be expanded in bigger areas and for different types of buildings i.e. commercial, residential. Moreover, deploying this method in the long term will result in huge paybacks. A zone-level HVAC system can achieve a huge amount of energy as it can operate proportionally to the population of each zone. Moreover, as each MAC address is unique to its user, each user can have his own profile with his preferable temperature and other settings.

In terms of localization, a smart phone software that can send its indoor location which determined through the neural network can be used in applications that seeking indoor localization i.e. tracking miners in mines, tracking employees in the office.

For future, it is recommended to do this experiment in a bigger area (a whole floor or building) with higher occupancy to further investigate the accuracy of this method. Also as our experiment was conducted at February month which was not so cold, the temperature difference with the set point was not a lot. Therefore, for future, this experiment can be done in the colder month of the year to check the energy consumption reduction. Moreover, the space usage can be investigated to see how well and by how many occupants the spaces (i.e. library, classrooms) are being utilized.

In summary, the wifi-based occupancy driven ventilation can replace the fixed ones commonly used to further assess the effect of occupancy patterns on building energy consumption.

Appendix

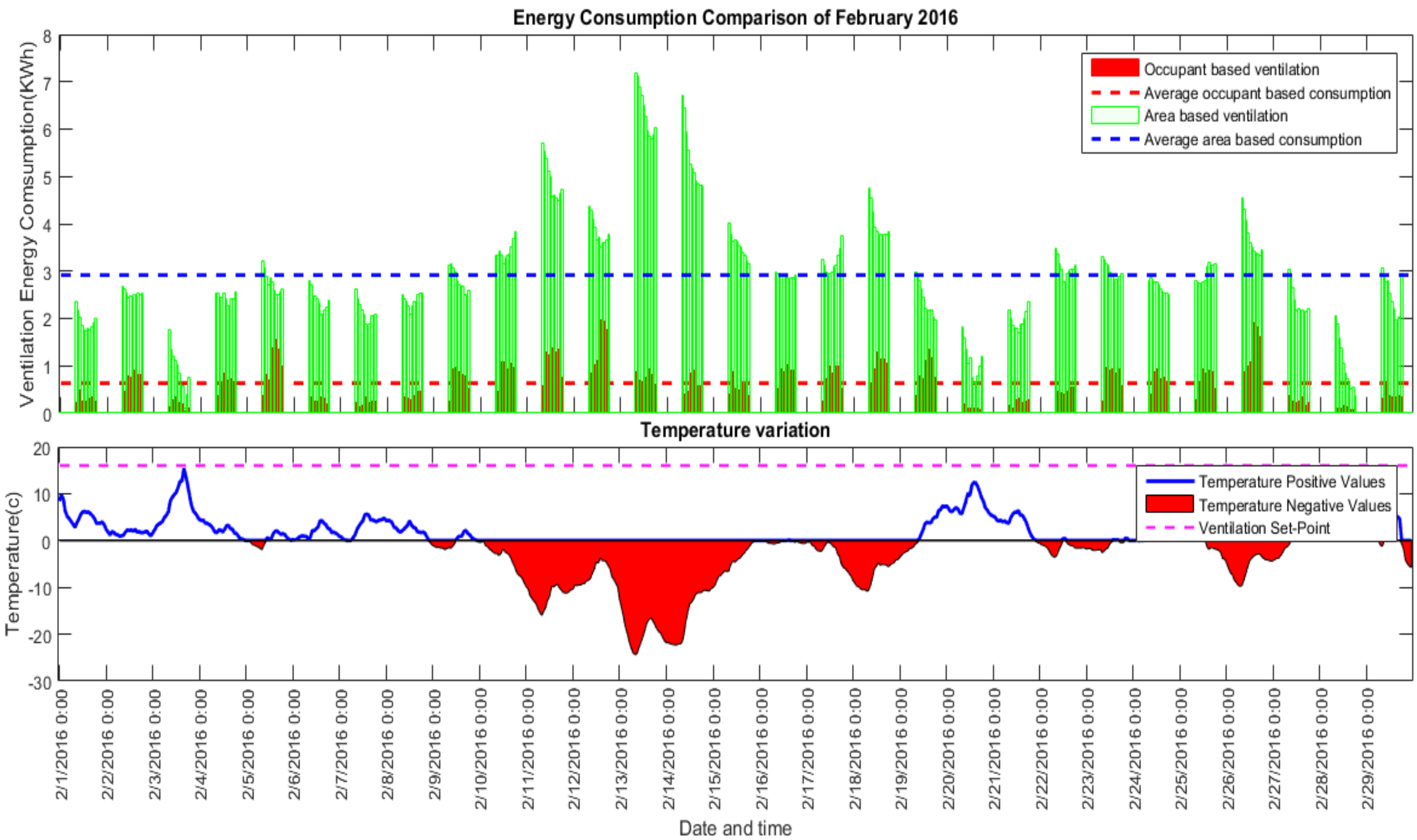

Figure 34. Energy consumption comparison of ventilation system (12Hr), February 2016

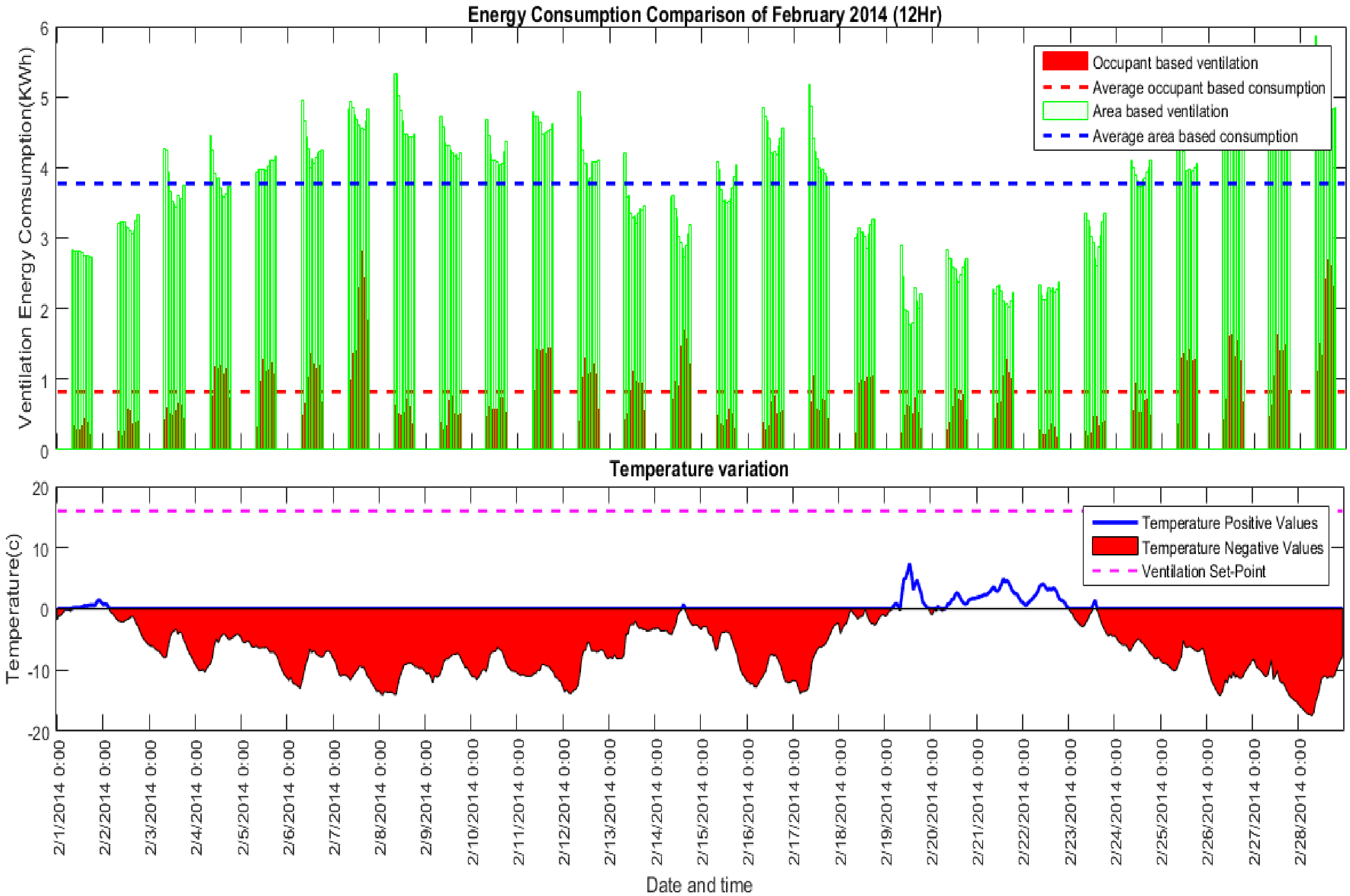

Figure 35. Energy consumption comparison of ventilation system (12Hr), February 2014

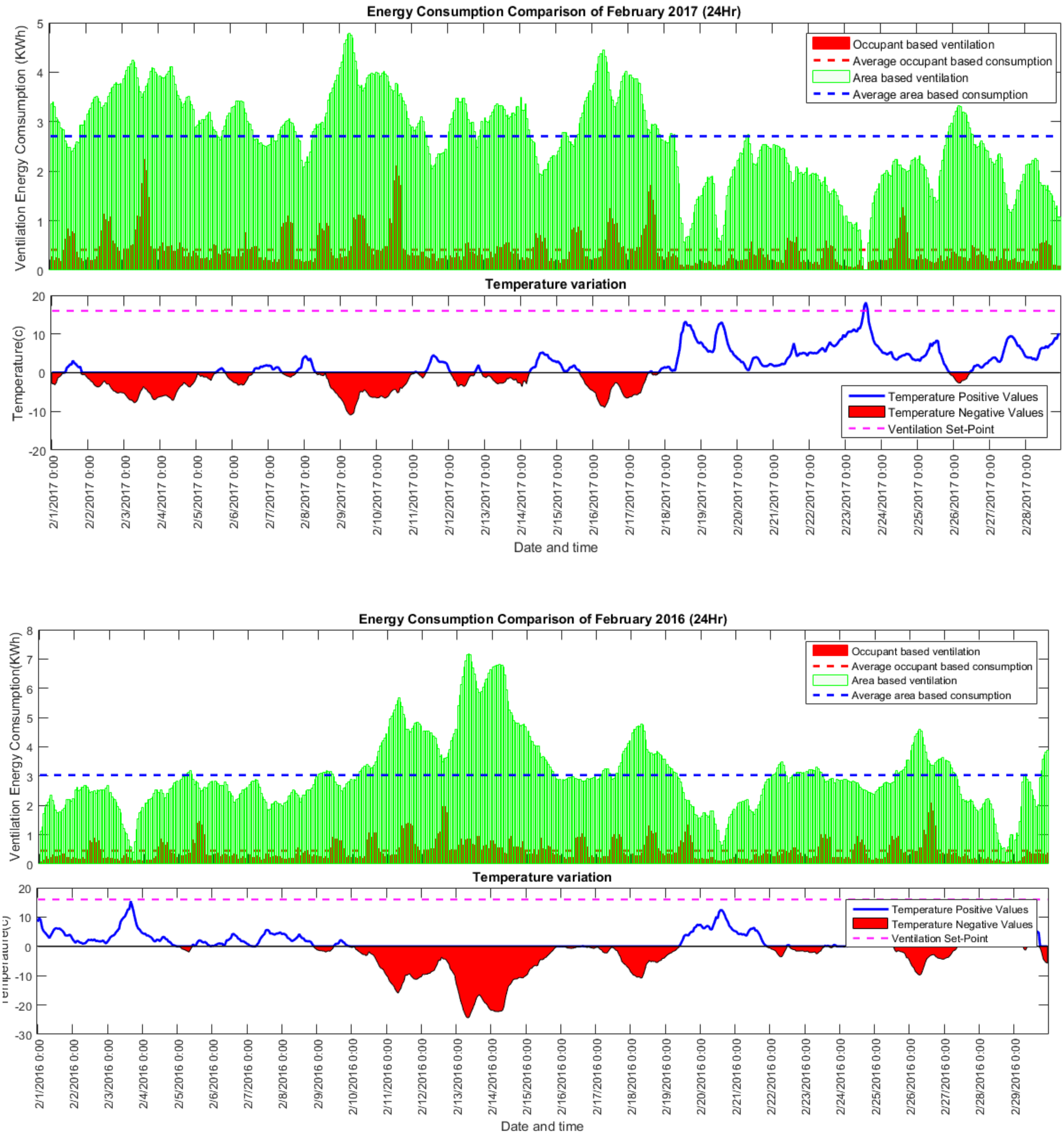

Figure 36. . Energy consumption comparison of ventilation system (24Hr), February 2017, 2016

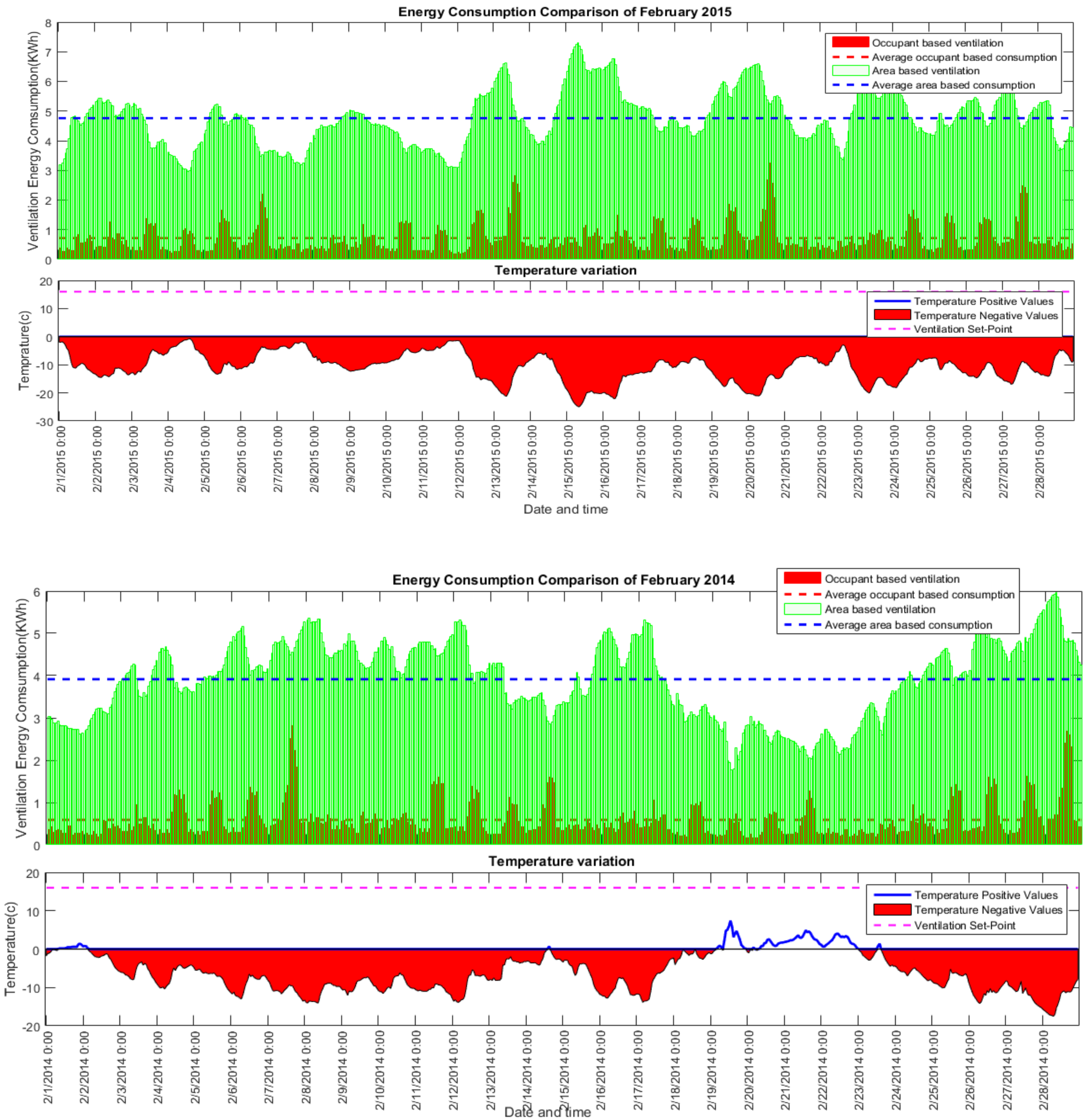

Figure 37. Energy consumption comparison of ventilation system (24Hr), February 2015, 2014

References

[1] *NRC, Natural Resources of Canada. "Energy fact book 2016-2017."*,2016.

[2] Z. Yu, B. C. M. Fung, F. Haghighat, H. Yoshino, and E. Morofsky, "A systematic procedure to study the influence of occupant behavior on building energy consumption," *Energy Build.*, vol. 43, no. 6, pp. 1409–1417, 2011.

[3] National Institute of Standards and Technology 'Measurement science roadmap for net-zero energy buildings workshop summary report', NIST Building and Fire Research Laboratory, NIST Technical Note 1660, 2010, http://www.energetics.com/news/Documents/NetZeroEnergyBuildings.pdf .

[4] "Iowa Energy Center. Lower energy bills with a set-back thermostat." http://www.energy.iastate.edu/news/pr/energysavingideas/setbacktherm.html.

[5] V. L. Erickson and A. E. Cerpa, "Occupancy based demand response HVAC control strategy," *Proc. 2nd ACM Work. Embed. Sens. Syst. Energy-Efficiency Build. - BuildSys '10*, p. 7, 2010.

[6] Agarwal, Y., Balaji, B., Dutta, S., Gupta, R. and Weng, T,'Duty-cycling buildings aggressively: the next frontier in HVAC control', Proceedings of ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN),2011.

[7] V. Erickson, M. Carreira-Perpinan, A. Cerpa, "OBSERVE: occupancy-based system for efficient reduction of HVAC energy",Information

Processing in Sensor Networks (IPSN), 10th International Conference on, pp. 258–269,2011.

[8] ASHRAE, *ANSI/ASHRAE Standard 62.1-2013 ,"Ventilation for Acceptable Indoor Air Quality"*.Atlanta, GA, American Society of Heating, Refrigerating, and Air Conditioning Engineers, Inc. 2013.

[9] Zhang, J. Lutes, R. G., Liu, G., Brambley, M.R. "Energy savings for occupancy-based control (OBC) of variable-air-volume (VAV) systems". PNNL-22072. Pacific Northwest National Laboratory. U.S. Department of Energy, 2013.

[10] Dougan, D.S., Damiano, L. "CO2-based demand control ventilation: do risks outweigh potential rewards?", ASHRAE Journal 46 (10). 47-53. 2004.

[11] Nassif, N. A robust CO2-based demand-controlled ventilation control strategy for multi-zone HVAC systems. Energy and Buildings, 45, 72-81,2012.

[12] Warren, B.F. and Harper. N.C., "Demand controlled ventilation by room CO2 concentration. A comparison of simulated energy savings in an auditorium space", Energy Build, Vol. 17, pp. 87-96, 1991.

[13] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, and T. Weng, "Occupancy-driven energy management for smart building automation," in Proc. ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building, 2010, pp. 1-6.

[14] S. Li, N. Li, B. Becerik-gerber, and G. Calis, "Rfid-Based Occupancy Detection Solution for Optimizing Hvac Energy Consumption," *Proc. 28th ISARC*, no. February 2016, pp. 587–592, 2011.

[15] Lee H., Jae S.C., Elmasri R., "A conflict resolution architecture for the comfort of occupants in intelligent office", IET, pp. 3-4, 2008

[16] V. Erickson, S. Achleitner, and A. Cerpa, "POEM: power-efficient occupancy-based energy management system," in Proc. IEEE Information Processing in Sensor Networks (IPSN), Apr 2013, pp. 203- 216.

[17] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, and Y. Agarwal, “Sentinel: occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings,” *Proc. 11th ACM Conf. Embed. Networked Sens. Syst.*, p. 17, 2013.

[18] D. Sánchez-Rodríguez, P. Hernández-Morera, J. Quinteiro, and I. Alonso-González, “A Low Complexity System Based on MultipleWeighted Decision Trees for Indoor Localization,” *Sensors*, vol. 15, no. 6, pp. 14809–14829, 2015.

[19] W. Xi *et al.*, “Electronic frog eye: Counting crowd using WiFi,” in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, 2014, pp. 361–369.

[20] H. Hong, C. Luo, and M. C. Chan, “SocialProbe: Understanding social interaction through passive WiFi monitoring,” *ACM Int. Conf. Proceeding Ser.*, vol. 28–Novembe, pp. 94–103, 2016.

[21] E. Vattapparamban, 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.

[22] K. Christensen *et al.*, “Using existing network infrastructure to estimate building occupancy and control plugged-in devices in user workspaces,” *Int. J. Commun. Networks Distrib. Syst. J. Commun. Networks Distrib. Syst.*, vol. 12, no. 1, pp. 4–29, 2014.

[23] K. Ehrhardt-Martinez and A. John, “People-centered initiatives for increasing energy savings,” 2010.

[24] O. T. Masoso and L. J. Grobler, “The dark side of occupants’ behaviour on building energy use,” *Energy Build.*, vol. 42, no. 2, pp. 173–177, 2010.

[25] T. A. Nguyen and M. Aiello, “Energy intelligent buildings based on user activity: A survey,” *Energy Build.*, vol. 56, pp. 244–257, 2013.

[26] H. Ben and K. Steemers, “Energy retrofit and occupant behaviour in protected housing: A case study of the Brunswick Centre in London,” *Energy Build.*, vol. 80, pp. 120–130, 2014.

[27] X. Ren, D. Yan, and C. Wang, "Air-conditioning usage conditional probability model for residential buildings," Building and environment, vol. 81, pp. 172-182, 2014

[28] P. Warren and L. Parkins, "Window-opening behaviour in office buildings," Building services engineering research and technology, vol. 5, pp. 89-101, 1984

[29] Y. Zhang and P. Barrett, "Factors influencing occupants’ blind-control behaviour in a naturally ventilated office building," Building and environment, vol. 54, pp. 137-147, 2012

[30] M. Foster and T. Oreszczyn, "Occupant control of passive systems: the use of Venetian blinds," Building and environment, vol. 36, pp. 149-155, 2001

[31] J. Ouyang, K. Hokao, Energy-saving potential by improving occupants’ behavior in urban residential sector in Hangzhou City, China, Energy and Buildings 41 (7) , 711–720 2009.

[32] M. Lopes, C. Antunes, and N. Martins, “Energy behaviours as promoters of energy efficiency: A 21st century review,”Journal of *Renewable and Sustainable Energy*, 2012.

[33] Y. Kim, Z. Charbiwala, A. Singhania, T. Schmid, M. B. Srivastava, SPORTLIGHT: Personal Natural Resource Consumption Profiler, HotEmNets, 2008

[34] M. Mozer, "The neural network house: an environment that adapts to its inhabitants", in: AAAI Spring Symposium on Intelligent Environments, pp. 110–114,1998.

[35] [51]Bode Project,2011,Webiste http://www.andes.ucmerced.edu/research/occupancy.

[36] J. Torres-Solis, T. Falk, and T. Chau, "A review of indoor localization technologies: towards navigational assistance for topographical disorientation," *Ambient Intell.*, 2010.

[37] H. Mehmood and N. Tripathi, "Indoor positioning system using artificial neural network," *Journal of Computer Science6 (10): 1206-1212, 2010*.

[38] Y. Kim, H. Shin, Y. Chon, and H. Cha, "Smartphone-based Wi-Fi tracking system exploiting the RSS peak to overcome the RSS variance problem," *Pervasive Mob. Comput.*, 2013.

[39] Z. Chen, H. Zou, H. Jiang, Q. Zhu, Y. Soh, and L. Xie, "Fusion of WiFi, smartphone sensors and landmarks using the Kalman filter for indoor localization," *Sensors*, 2015.

[40] Y. Chen, J. Chiang, H. Chu, and P. Huang, "Sensor-assisted wi-fi indoor location system for adapting to environmental dynamics," *Proc. 8th*, 2005.

[41] C. Feng, W. Au, S. Valaee, and Z. Tan, "Received signal strength based indoor positioning using compressive sensing," *IEEE Trans. Mob. Comput.*, pp. 1–12, 2011.

[42] Maniccia, D., Rutledge, B., Rea, M. S., & Morrow, W. (1999). Occupant use of manual lighting controls in private offices. Journal of the Illuminating Engineering Society, 28, 42–56.

[43] J. Brooks, S. Kumar, S. Goyal, R. Subramany, and P. Barooah, “Energy-efficient control of under-actuated HVAC zones in commercial buildings,” *Energy Build.*, vol. 93, pp. 160–168, 2015.

[44] G. R. Newsham and B. J. Birt, “Building-level occupancy data to improve ARIMA-based electricity use forecasts,” *Proc. 2nd ACM Work. Embed. Sens. Syst. Energy-Efficiency Build. - BuildSys ’10*, p. 13, 2010.

[45] D. Floyd, D. Parker, and J. Sherwin, “Measured Field Performance and Energy Savings of Occupancy Sensors: Three Case Studies.,” 1996.

[46] F. Haldi and D. Robinson, "Stochastic / probabilistic modelling of multiple adaptive processes: some subtle complexities," presented at the eSim Quebec City, QC, May 20-21, 2008.

[47] A. Mahdavi, A. Mohammadi, E. Kabir, and L. Lambeva, "Occupants' operation of lighting and shading systems in office buildings," Journal of Building Performance Simulation, vol. 1, pp. 57-65, 2008.

[48] D. Yan *et al.*, “Occupant behavior modeling for building performance simulation: Current state and future challenges,” *Energy Build.*, vol. 107, pp. 264–278, 2015.

[49] A. Rahman, P. Devi, and S. Partiwi, “Occupant movement tracking in high-rise building evacuation planning using RFID technology and SMS gateway,” *Netw. Ergon.*, 2012.

[50] W. Wang, R. Joshi, A. Kulkarni, W. K. Leong, and B. Leong, “Feasibility study of mobile phone WiFi detection in aerial search and rescue

operations," in *Proceedings of the 4th Asia-Pacific Workshop on Systems - APSys '13*, 2013, pp. 1–6.

[51] E. Verbree, S. Zlatanova, and K. van Winden, "To localise or to be localised with WiFi in the Hubei museum?" 2013.

[52] Osa, V., Matamales, J., Monserrat, J., & López, J. Localization in wireless networks: The potential of triangulation techniques. Wireless Personal Communications, 68(4), 1525-1538,2013.

[53] Kushki, A., Plataniotis, K. and Venetsanopoulos, A. 'Kernel-based positioning in wireless local area networks', IEEE Transactions on Mobile Computing, Vol. 6, No. 6, pp.689–705, 2007.

[54] Khoury, H. M., & Kamat, V. R.. Evaluation of position tracking technologies for user localization in indoor construction environments. Automation in Construction, 18(4), 444-457.doi:10.1016/j.autcon.2008.10.011,2009.

[55] I. Goodfellow. and Y. Bengio. and A. Courville, "*Deep LearningBook*" MIT Press, 2016.

[56] R. Melfi, B. Rosenblum, B. Nordman, and K. Christensen, "Measuring building occupancy using existing network infrastructure BT - 2011 International Green Computing Conference, IGCC 2011, July 25, 2011 - July 28, 2011," 2011.

[57] Liu, G., J. Zhang, and A. Dasu." Review of Literature on Terminal Box Control, Occupancy Sensing Technology and Multi-zone Demand Control Ventilation (DCV)". PNNL-21281, Pacific Northwest National Laboratory, Richland, WA,2012.

[58] Government of Canada, “Hourly data report,” Toronto,2017.http://climate.weather.gc.ca/historical_data/search_historic_data_e.html

[59] H. Mehmood, N. K. Tripathi, and T. Tipdecho, “Indoor positioning system using artificial neural network,” *J. Comput. Sci.*, vol. 6, no. 10, pp. 1219–1225, 2010.

[60] N. M. Nawi, W. H. Atomi, and M. Z. Rehman, “The Effect of Data Pre-processing on Optimized Training of Artificial Neural Networks,” *Procedia Technol.*, vol. 11, pp. 32–39, 2013.

[61] Lahiri, S. K., & Ghanta, K. C. (2010). Artificial neural network model with parameter tuning assisted by genetic algorithm technique: Study of critical velocity of slurry flow in pipeline. Asia- Pacific Journal of Chemical Engineering, 5(5), 763–777.

[62] Popleteev, A., & Engel, T. (2014). Device-Free Indoor Localization Based on Ambient FM Radio Signals. International Journal of Ambient Computing and Intelligence, 6(1), 35–44. http://doi.org/10.4018/ijaci.2014010103