Understanding Occupancy Patterns in a Commercial Space

Kin Sum Liu[1], Sirajum Munir[2], Shan Lin[1], and Charles Shelton[2]

[1]Stony Brook University, Stony Brook, NY
[2]Bosch Research and Technology Center, Pittsburgh, PA

1 Introduction

Heating, ventilation, and air conditioning (HVAC) is a major source of energy consumption in the US. In 2006, approximately 35% of energy in the US was used for HVAC [1]. Usually building operators use a static schedule for controlling HVAC systems without having a deeper understanding of how many people use the building at different times of the day. In addition, HVAC systems operate by assuming maximum occupancy in each room, which leads to a significant energy waste, e.g., an HVAC system providing ventilation for 30 people when there are only 10 people in a room [9]. Understanding how many people use different rooms at different times of the day is very crucial for achieving building energy efficiency and providing occupant comfort.

There have been several attempts to collect long term occupancy patterns from office buildings, e.g., Mitsubishi's Electronic Research Lab (MERL) dataset [14] and Colorado School of Mines (CSMBB) dataset [4]. However, PIR motion detectors are used in these projects for sensing occupancy and hence these datasets only reflect whether a room is occupied or not without revealing the actual person count. Similar datasets are also collected from households [2][11]. Only a few datasets from commercial spaces capture actual person count, and even in these cases, the data collection period was very limited, e.g., in [3] only 5 days of occupancy data is collected as ground truth. We are the first to collect long term (9 months) occupancy data (people count) from an 11,000 square foot commercial office (second floor of Bosch Research and Technology Center Pittsburgh). In this paper, we outline some statistical results from our empirical study and explore their implications for saving energy.

The findings from the empirical study are about understanding and predicting occupancy patterns (1, 2, 3, 4), understanding the usage of a common office space (5, 6) and conference rooms (7, 8), and weekend behavior (9) as described below.

1. The daily occupancy pattern of the common office space is regular. However, the percentage change of the peak (maximum occupancy) varies from 0% to 730% with an average of 18.4%.

2. The daily occupancy pattern in conference rooms is very irregular.

3. Simple solution leveraging occupancy count from the previous day for predicting future occupants suffers a percentage error from 0% to 1100% with an average of 30.1% for the common office space. It is much worse for the conference rooms, e.g., in one conference room (Warhol) it varies from 0% to 2100% with an average of 104%.

4. The number of people in the initial 3 minutes is a useful indicator for predicting number of participants in a meeting. Such a simple predictor achieves 2.6 people RMSE error on average.

5. Arrival time of the earliest person is at 7:25 AM on average, which is 1.5 hours before the arrival of the earliest group of people, which is 10 people in this analysis.

6. Departure time of the latest person is 8:43 PM on average which is 3 hours after the latest group.

7. In conference rooms, the average and median number of participants in meetings are 3.85 and 3, respectively. We find that in 99% cases, the number of participants does not reach the maximum occupancy capacity of the conference rooms. In fact, the number of participants does not exceed half of the room capacity in 95% cases. The average and median durations of meetings are 63 minutes and 48 minutes, respectively.

8. The utilization of conference rooms varies from different times of the day with a mean of 55.6% between work hours (9 AM to 6 PM), which means conference rooms are unused in 44.4% time during work hours.

9. The office space is not completely unused during weekends. The chance of occupancy in a day for the common office space is 57.9% and the utilization of conference rooms is 4.8% on average during weekends.

We published some of these findings in a previous work briefly [5]. In this paper, we describe the findings in more details. The rest of the paper is organized as follows. In Section 2, we describe the deployment and data collection procedure. Then, we describe the findings in detail in Section 3. Finally, we discuss about the findings, their implications for saving energy, and our future work in Section 4.

2 Deployment and Data Collection

There are several solutions for occupancy estimation using IR-array sensors [8], ultrasonic sensors [13], and RGB cameras [12]. We use a depth sensor (Kinect for XBox One) for counting number of people in a room. The depth sensor is mounted on the ceiling looking downwards near to a doorway as it accurately counts the number of people entering and exiting through the door. The solution is called FORK (Fine grained Occupancy estimatoR using Kinect) and the detailed algorithm is described in [9, 10]. We have deployed five instances of FORK at a Bosch office to cover the *common office space* (requires monitoring at two entrances: Main Gate and Remote Gate), two *conference rooms* (Warhol and Clemente), and a lab. Warhol and Clemente conference rooms are the most used and largest conference rooms of this office that can accommodate 25 and 20 people, respectively. The floor plan of the office is shown at Figure 1. Four FORK units were deployed on August 24th 2015. The Remote Gate was not used at that time often, but later a group migrated to the other side of the office and hence the Remote Gate unit was deployed on January 27th 2016. After collecting over 9 months of data, we stopped the data collection on June 7th 2016. During this period, there were 64,925, 18,477, 9,938 entrance and exit events in the common office space (Main Gate and Remote Gate), Warhol, and Clemente conference rooms, respectively. The entire dataset is available at [6].

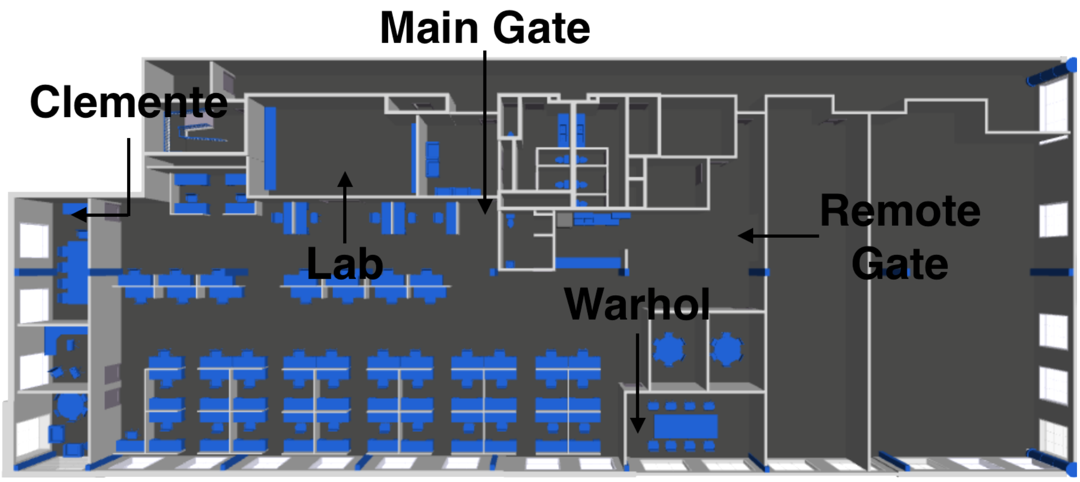


Figure 1: Floor plan of the office

3 Findings in Detail

In this section, we describe the findings from the long term empirical study in detail.

3.1 Occupancy patterns in the common office space: Occupancy count of the common office space from April 2016 is shown in Figure 2. It shows that the shape of waveform is consistent across days and weeks in the sense that the influx and outflux of people happen at similar times. However, the magnitude (maximum people count in a day) varies, which is shown in Figure 3.

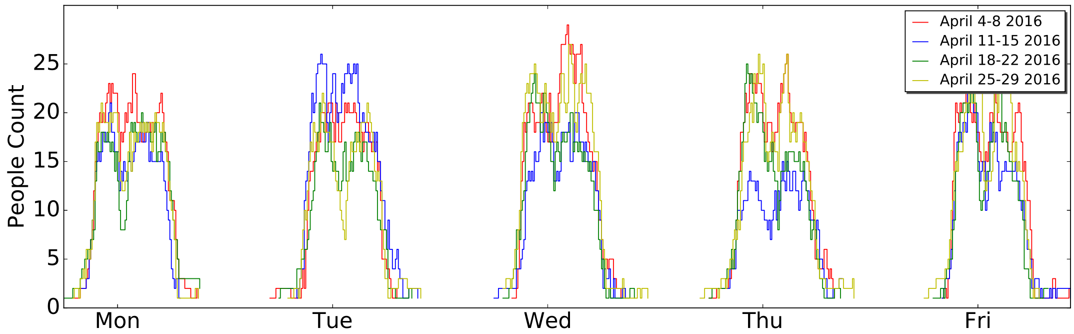


Figure 2: People count in common office space in April 2016

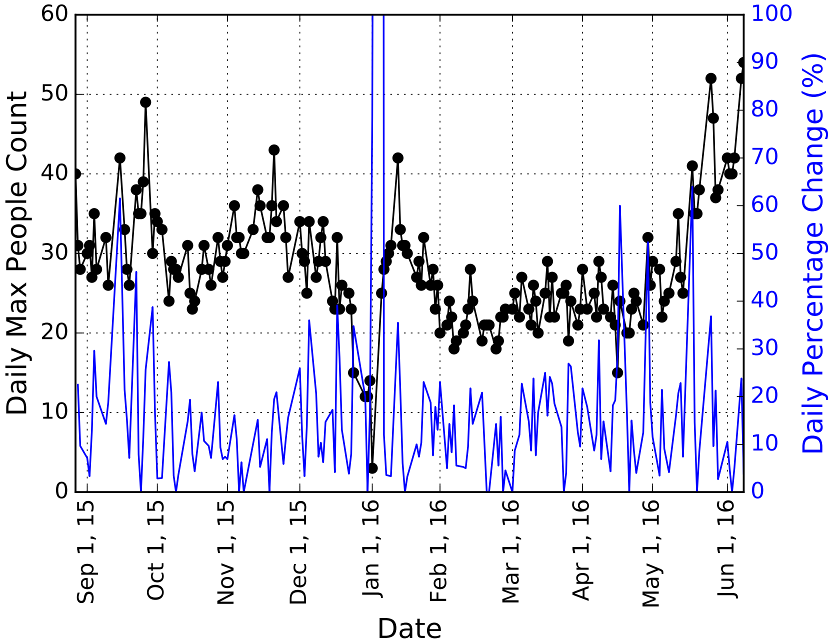


Figure 3: Variation of maximum people count in the common office space

The black curve shows a large variation of daily maximum people count with an average day-to-day change of 4.1 persons. The percentage change is 18.4% on average with a maximum of 730% on December 31st 2015 (a cut off portion is shown in the figure) and a minimum of 0%. The reason for this variation is interns joining and leaving, joining of new employees, arrival of visitors, arrival of interview candidates, employees travelling and on vacation, and different teams migrating to different parts of the building.

3.2 Occupancy patterns in conference rooms: Occupancy pattern of a conference room (Warhol) for the same month (April) is shown in Figure 4. It shows that occupancy patterns are not consistent even across days. The timing, frequency, size and duration of meetings vary significantly in different days.

3.3 Prediction of future occupancy: One simple way to predict occupancy pattern is by averaging the historical occupancy data of previous days. We define a time slot as a 15 minute interval and predict the number of people in a particular slot by averaging the occupancy count of the same slot in the previous n days (Figure 5(a)) and same weekdays in the previous n weeks (Figure 5(b)). We vary n from to 1 to 5 in the x-axis of both figures. We show both

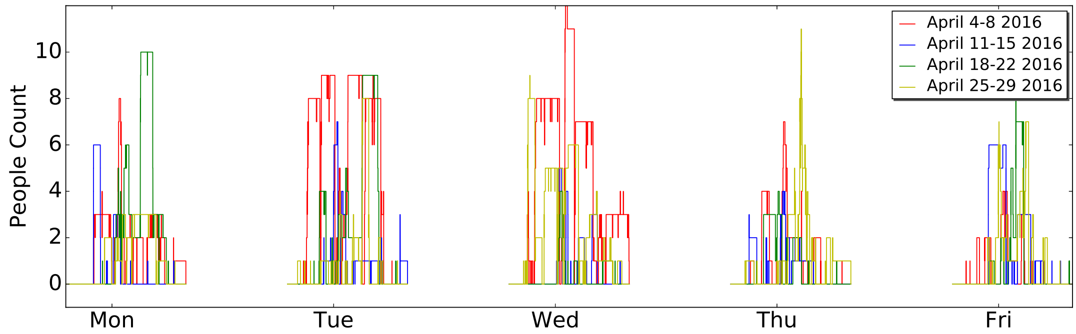
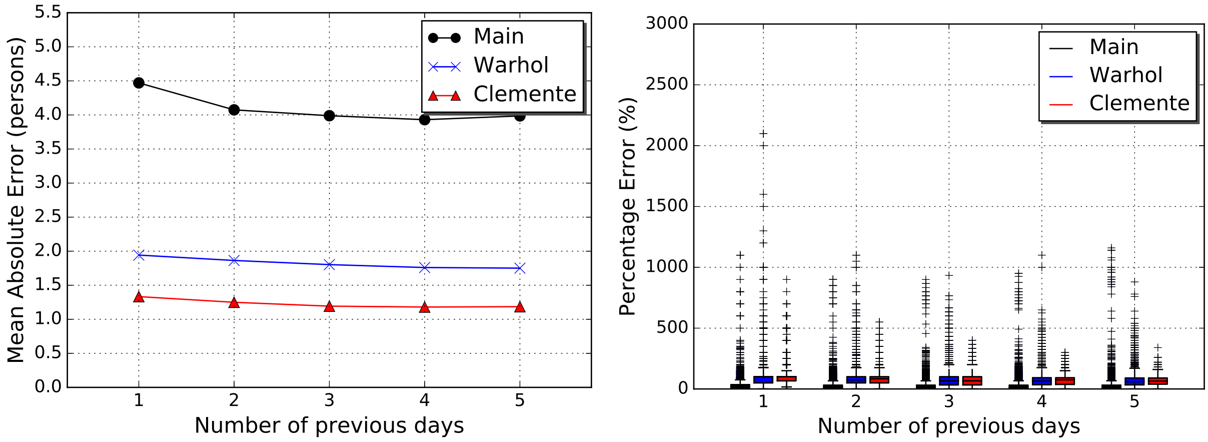
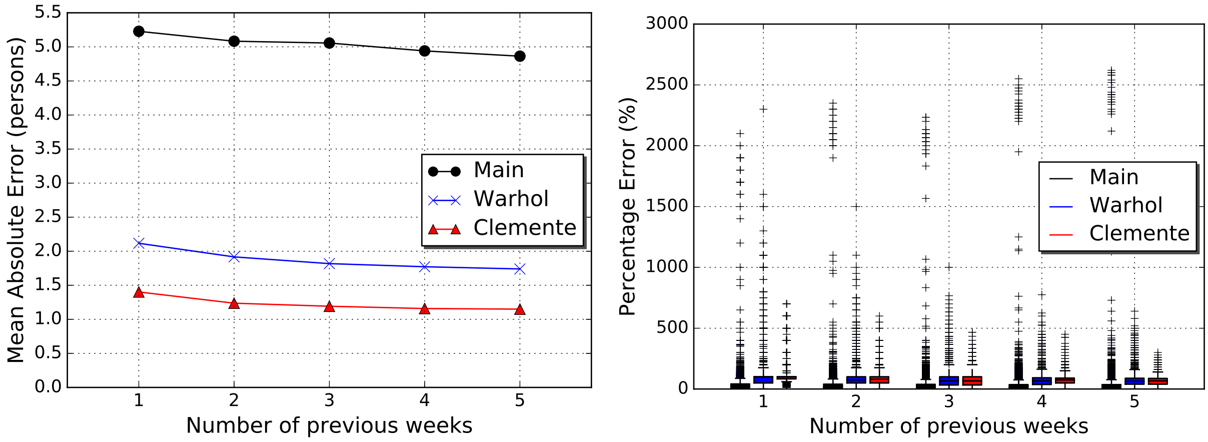


Figure 4: People count in a conference room (Warhol) in April 2016



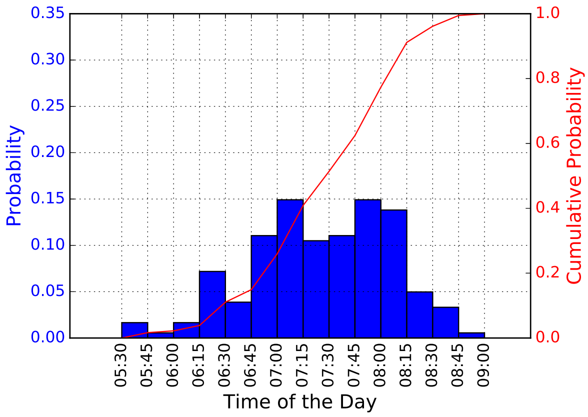
(a) Using previous days for prediction



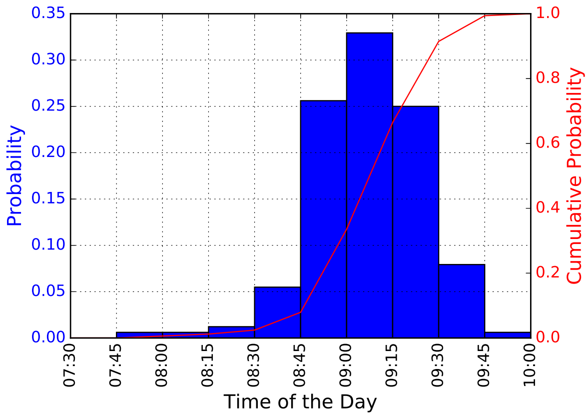
(b) Using previous weeks for prediction

Figure 5: Prediction errors using previous data points to predict future occupancy count from 9:00 to 18:00

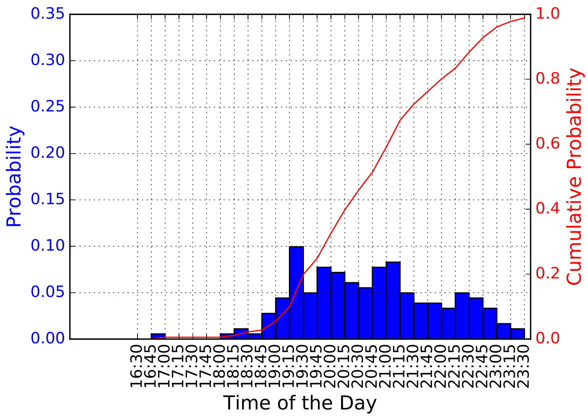
Mean Absolute Error (offset from the actual count) and box plot for Percentage Error (percentage offset from the actual count) in the sub-figures. The figures show that the prediction for the count of common office space suffer a large error and the error for conference rooms is even more severe. Using the previous day for prediction, the means of percentage errors are 30.1%, 104% and 97.6% for common office space, Warhol and Clemente, respectively. We exclude 0.2%, 48.2%, and 61.5% instances when the actual count is 0 and hence percentage error becomes undefined. If the previous week's data is used, the means of percentage errors are increased to 38.5%, 112% and 98.1%,



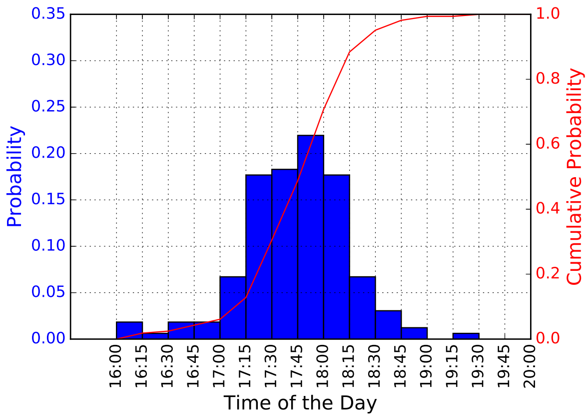
(a) Earliest person's arrival



(b) Earliest Group's arrival



(c) Latest person's departure



(d) Latest Group's departure

Figure 6: Distribution of arrival and departure times in the common office space

respectively. We exclude 0.2%, 48.4%, and 61.3% instances when the actual count is 0 and percentage error is undefined.

3.4 Prediction of number of participants: Figure 7 shows how error varies when number of participants at different times from the beginning of a meeting is used as a direct predictor of the maximum number of participants in the meeting. It shows that after 1 minute of when a meeting starts, the Root Mean Square Error reaches below 2 and 4 persons from the actual number of participants with a Mean Percentage Error at 29% and 35% for Clemente and Warhol, respectively. If this time interval is increased to 3 minutes, RMSE is reduced to 1.6 and 3.6, while the percentage error is 21% and 26% for Clemente and Warhol, respectively.

3.5 Arrival time: The arrival time of the earliest person ranges between [5:30, 9:00], which is shown in in Figure 6(a). For 50% of the days, we observe people arriving as early as 7:30. We define a group consisting of at least 10 people. The arrival time of the earliest group is shown in Figure 6(b), which shows that the range is much smaller spanning [7:45, 10:00] and it happens between [8:30, 9:30] usually. On average, the earliest person arrives 1.5 hours before the earliest group.

3.6 Departure Time: We show the departure time of the latest person in Figure 6(c), which shows that the departure time ranges between [16:45, 23:30]. The outlier 16:45 happened on December 31st 2015. If we exclude this data point, the span is still across 4 hours. The departure time of the latest group is shown in Figure 6(d), which shows that

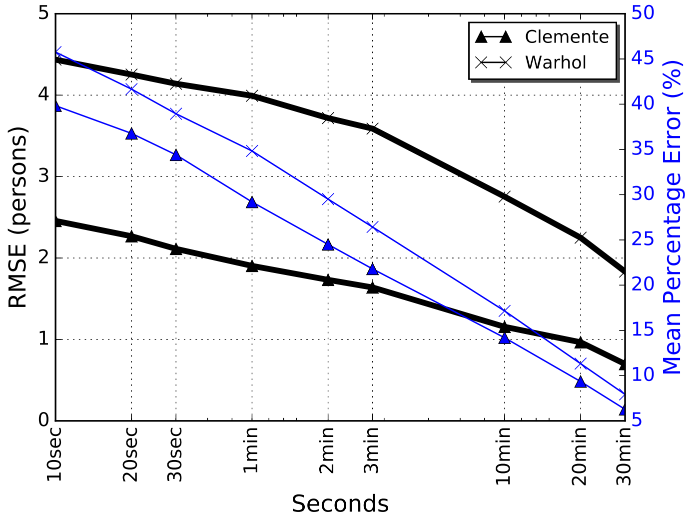


Figure 7: Prediction error on size of meetings based on number of people at the beginning of meetings

the range is smaller spanning between [16:00, 19:30] and this happens between [17:00, 18:30] usually. On average, the latest person leaves almost 3 hours after the latest group.

3.7 Duration and number of occupants in meetings: The main usage of conference rooms is for meetings. We define a meeting as an occupancy of more than one person in the room for more than 1 minute. The distribution of duration and number of participants in meetings in the conference rooms (Clemente and Warhol) are shown in Figures 8 and 9, respectively. Average durations of a meeting in Clemente and Warhol are 62.2 minutes and 65.3 minutes, respectively. Number of participants in a meeting in Clemente and Warhol are 3.4 and 4.3 people on average. Both metrics have long tails that indicates large or long meetings happen infrequently but they exist. Clemente's and Warhol's capacity is 20 and 25 people respectively. For both conference rooms, 99% of meetings have not reached the maximum capacity. Also, 95% of meetings have the number of participants equal to or below 7 and 12, respectively. Therefore, for most of the times, they are occupied by less than 50% of maximum capacity. The outlier happened once on November 24th 2015 when Warhol hosted 29 people for a Thanksgiving party.

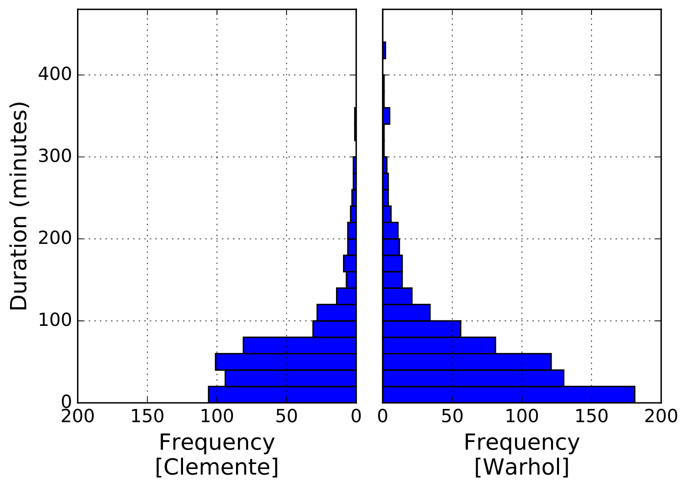


Figure 8: Distribution of duration of meetings.

3.8 Utilization of conference rooms: Figure 10 shows the utilization rate of both conference rooms with respect to

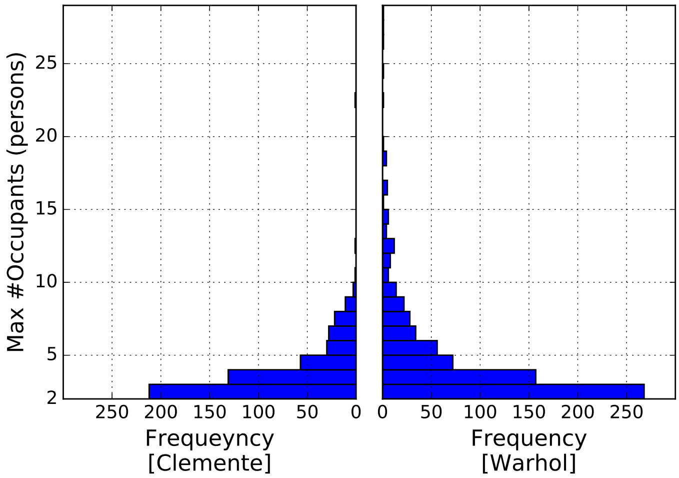


Figure 9: Distribution of number of occupants in meetings.

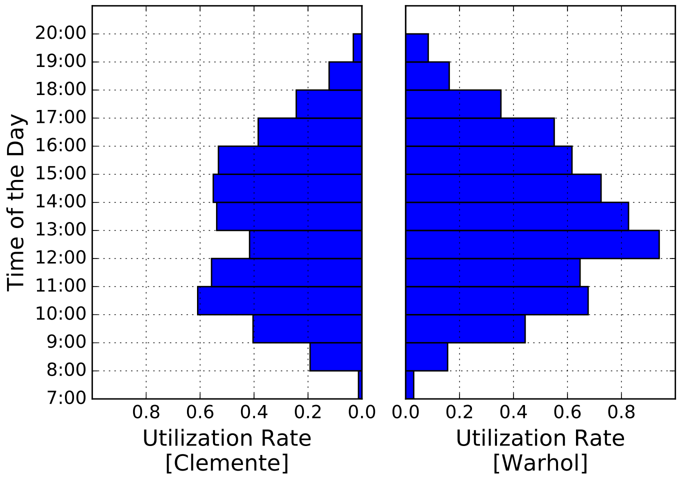


Figure 10: Hourly utilization rate of two conference rooms from 7:00 to 20:00.

different times of the day over the whole observation period. It is utilized if there is at least one meeting in that hour. The average utilization rate of Clemente and Warhol for between 9:00 to 18:00 is 47.0% and 64.2% respectively. Therefore, the average utilization between work hours (9 AM to 6 PM) is 55.6%. In the early morning before 9:00 and late afternoon after 18:00, the chance of each room being used is less than 20%. The two conference rooms also have different patterns of usage. For both conference rooms, the utilization rate varies greatly throughout the day. First, Clemente is used less frequently. The utilization does not exceed 65% for any time of day. There is also a drop in utilization rate at noon because employees take their lunch break. On the contrary, Warhol is used for around 70% from 10:00 to 16:00. The peak appears at noon because some employees have their lunch in Warhol at that time.

3.9 Occupancy patterns during weekends: Throughout the data collection period, there is a chance of 57.9% that at least one person comes to the common office space on a day during weekends. On average, the office is used for 187 minutes with a standard deviation of 212 minutes. The conference rooms are rarely used. The chance of usage is 2.78% and 6.85% for Clemente and Warhol, respectively. The average utilization is 4.8%.

4 Discussions and Future Works

In addition to detecting humans, it is useful to sense additional contextual information regarding them for an efficient HVAC control. In our recent work [7], we detect objects that are carried inside and outside of a room in a building, e.g., backpacks, boxes, laptops, phones, umbrellas, cups, and guns. In addition to improving safety and security of the occupants, object detection can be useful for controlling HVAC systems to save energy. For example, in a commercial space, if someone is leaving his office with a backpack in the afternoon, it may mean he is leaving for the day, whereas if he is leaving with a laptop, it may mean he is going out for a meeting, and leaving empty handed may mean that he is leaving for a restroom/lunch break. If someone enters in a conference room with a phone at his hand/ear, he is probably going there to make a phone call, which may not last as long as a regular office meeting. An HVAC system can potentially employ a better control strategy with such additional information.

Our empirical study suggests several findings for more energy efficient HVAC control. We see that the variance of the earliest arrival and latest departure time for an individual is much larger than that of a group behavior. Hence, if we sacrifice the comfort of the first and last few people, that will reduce energy consumption substantially. In that way, instead of running the HVAC system at a full rate from 7:00 to 23:30, we can run it at a lighter load between 6:30 to 8:30 and between 18:30 to 23:30. Occupancy count of the common space shows a periodic behavior with a variation of amplitude (maximum occupancy). Hence, an adaptive solution is useful for predicting occupancy in the common space. However, occupancy prediction is very difficult for conference rooms. We see that occupancy in the first 3 minutes is a good indicator of the number of participants in a meeting in conference rooms. Hence, a multistage HVAC system can be useful for saving energy. Another reason for using it is that conference rooms are not used for 44.4% time during office hours. Hence, instead of keeping the rooms warm/cool during the entire work day, the rooms can be put to a setback threshold and be heated/cooled based on demand. Also, we find that in 95% cases, the number of occupants does not exceed half of the capacity of conference rooms. Hence, a Variable Air Volume (VAV) HVAC system will save energy. In the future, we plan to use a simulation tool to determine the amount of energy savings possible using these techniques.

5 Acknowledgments

This work was supported, in part, by DOE grant DE-EE0007682 and NSF grant CNS 1536086. The opinions expressed here are those of the authors and do not necessarily reflect the views of the DOE.

References

[1] EIA - energy information administration. http://www.eia.doe.gov/.

[2] Christian Beckel, Wilhelm Kleiminger, Romano Cicchetti, Thorsten Staake, and Silvia Santini. The eco data set and the performance of non-intrusive load monitoring algorithms. In *Proceedings of the 1st ACM International Conference on Embedded Systems for Energy-Efficient Buildings (BuildSys)*, pages 80–89. ACM, November 2014.

[3] Varick L. Erickson, Miguel Á. Carreira-Perpiñán, and Alberto Cerpa. OBSERVE: occupancy-based system for efficient reduction of HVAC energy. In *IPSN*, 2011.

[4] W. A. Ho and J. W. Howard. Activity recognition in a dense sensor network. In *International Conference on Sensor Networks and Applications (SNA)*, 2009.

[5] K. S. Liu, S. Munir, J. Francis, C. Shelton, and S. Lin. Poster abstract: Long term occupancy estimation in a commercial space: an empirical study. In *IPSN*, 2017.

[6] K. S. Liu, E. V. Pinto, S. Munir, J. Francis, C. Shelton, M. Berges, and S. Lin. Poster: COD: a dataset of commercial building occupancy traces. In *BuildSys*, 2017.

[7] N. C. Mithun, S. Munir, K. Guo, and C. Shelton. ODDS: real-time object detection using depth sensors on embedded GPUs. In *IPSN*, 2018.

[8] H. Mohammadmoradi, S. Munir, O. Gnawali, and C. Shelton. Measuring people-flow through doorways using easy-to-install ir array sensors. In *DCOSS*, 2017.

[9] S. Munir, R. S. Arora, C. Hesling, J. Li, J. Francis, C. Shelton, C. Martin, A. Rowe, and M. Berges. Real-time fine grained occupancy estimation using depth sensors on ARM embedded platforms. In *RTAS*, 2017.

[10] S. Munir, L. Tran, J. Francis, C. Shelton, R. S. Arora, C. Hesling, M. Quntana, A. K. Prakash, A. Rowe, and M. Berges. Demo: FORK: Fine grained Occupancy estimatoR using Kinect on arm embedded platforms. In *BuildSys*, 2017.

[11] James Scott, AJ Bernheim Brush, John Krumm, Brian Meyers, Michael Hazas, Stephen Hodges, and Nicolas Villar. Preheat: controlling home heating using occupancy prediction. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 281–290. ACM, 2011.

[12] Jakub Segen and Sarma Pingali. A camera-based system for tracking people in real time. In *International Conference on Pattern Recognition*, volume 3, pages 63–67, Aug 1996.

[13] Oliver Shih and Anthony Rowe. Occupancy estimation using ultrasonic chirps. In *ICCPS*, 2015.

[14] C.R. Wren, Y.A. Ivanov, D. Leigh, and J. Westhues. The MERL motion detector dataset. In *Workshop on Massive Datasets (MD)*, pages 10–14, November 2007.