Poster Abstract: COD - A Dataset of Commercial Building Occupancy Traces

Kin Sum Liu
Stony Brook University
kiliu@cs.stonybrook.edu

Elvin Vindel Pinto
Calvin College
ejv56@students.calvin.edu

Sirajum Munir
Bosch Research and Technology Center, Pittsburgh, PA
sirajum.munir@us.bosch.com

Jonathan Francis
Bosch Research and Technology Center, Pittsburgh, PA
jon.francis@us.bosch.com

Charles Shelton
Bosch Research and Technology Center, Pittsburgh, PA
charles.shelton@us.bosch.com

Mario Berges
Carnegie Mellon University
marioberges@cmu.edu

Shan Lin
Stony Brook University
shan.x.lin@stonybrook.edu

ABSTRACT

In this poster we introduce the Commercial Occupancy Dataset (COD), a high-resolution long-term dataset of occupancy traces in a commercial office building spanning 9 months and covering zone-level occupancy for five different spaces containing more than 90,000 enter/exit events over this time period. Occupancy data in a building contains rich spatial-temporal information about the users and their usage of the space and facilities. However, obtaining accurate occupancy data is a very challenging task due to the limitation of existing sensing technologies. A novel depth-imaging based solution to estimate occupancy counts was deployed in five doorways of an office building to generate the dataset. The dataset has high resolutions in all three major dimensions: temporal, spatial and occupancy state. This allows applications such as building energy simulation, occupancy modeling and human-in-the-loop HVAC control which enhance energy efficiency and human comfort. Towards this objective, we provide a case study to demonstrate the utility of understanding occupancy.

ACM Reference Format:

Kin Sum Liu, Elvin Vindel Pinto, Sirajum Munir, Jonathan Francis, Charles Shelton, Mario Berges, and Shan Lin. 2017. Poster Abstract: COD - A Dataset of Commercial Building Occupancy Traces. In *Proceedings of 4th ACM Conference on Systems for Energy-Efficient Built Environments (BuildSys'17).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3137133.3141471

1 INTRODUCTION

Occupancy plays a significant role in the design and operation of buildings, especially as it relates to its effects on the energy consumed by facilities. Though research has shown that accurate occupancy estimation can increase the accuracy of building energy simulation models [6], reduce energy consumption [1] and support safety/security applications, most of these benefits are currently unobtainable due to the difficulties associated with collecting accurate occupancy count datasets. Moreover, because of the sensitivity of these purported benefits to the occupancy traces used, access to realistic occupancy data becomes critical. To this end, in this poster we present an open dataset of occupancy traces for a commercial office building in Pittsburgh, PA.

The Commercial Occupancy Dataset (COD), is a high-resolution long-term dataset of occupancy traces spanning 9 months and covering zone-level occupancy for five different spaces containing more than 90,000 enter/exit events over this time period. The rest of this paper is dedicated to explaining the characteristics of the dataset (including how it was collected), and providing illustrative examples of its potential usage by researchers and practitioners.

2 DATA COLLECTION

The data was collected using an occupancy estimation solution called FORK (Fine grained Occupancy estimatoR using Kinect) [5]. We have deployed five instances of FORK at a Bosch office in Pittsburgh (the floorplan is shown in Figure 1 with arrows pointing at deployment locations) to cover the common office spaces. The solution is based on a depth-imaging camera mounted on doorways with the camera positioned such that its field of view captures the area under the doorway. Though a total of five units were deployed (as shown in Figure 1 the lab measurements were erroneous, and the results of the gates were combined, producing a total of 3 spaces that were tracked.

The FORK algorithm [5] processes the depth images collected at close to 9 frames per second in real-time on an Odroid, and produces estimates for enter/exit events timestamped to match when people cross a pre-configured threshold representing the location of the doorway. The data contains over 9 months of data, during which time there were around 90,000 events in total. Each event triggered an update of the occupancy count inside the room. Thus, each entry in the dataset is a tuple of *(date, time, occupancy count).* Figure 2 shows a data snapshot for five weekdays in April 2016.

BuildSys'17, November 8–9, 2017, Delft, The Netherlands

ACM ISBN 978-1-4503-5544-5/17/11.
https://doi.org/10.1145/3137133.3141471

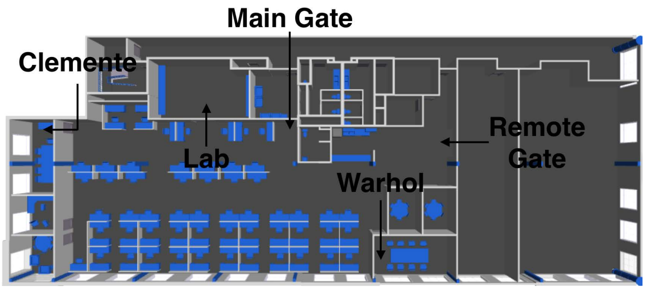

Figure 1: Floor plan of the office and location of sensors

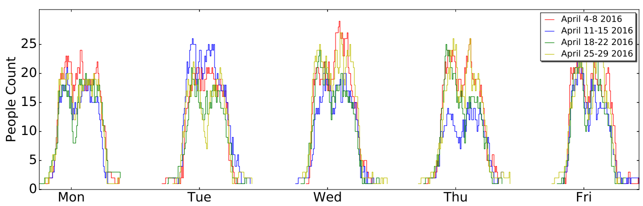

Figure 2: People count in common office space in April 2016

3 DATA CHARACTERISTICS

The dataset can be found at https://doi.org/10.5281/zenodo.996587 [4], and consists of five CSV files (one for each room) containing the occupancy traces for the rooms we monitored.

Duration: The dataset contains a total of 9 months. It covers scenarios with different temporal granularities including stochastic instances such as security guards patrolling and guests visiting, daily events like holidays.

Accuracy and robustness: The quality of our occupancy sensing system FORK is analyzed in [5] with more detail. Overall, the algorithm estimates occupancy with more than 95% accuracy, and can gracefully deal with complex images such as detecting multiple persons entering at the same time and distinguishing between human and non-human objects (such as people carrying equipment).

4 APPLICATIONS

Data module in simulation framework: Whole building simulation frameworks are very useful for understanding energy consumption in buildings. Our dataset can serve as a data module in carrying out long term and realistic simulations. For example, EnergyPlus [2] allows users to provide occupancy schedule for simulations. With our occupancy information and building model imported, the energy consumption from the simulation will draw a meaningful inference that applies to the real world.

Occupant behavior modeling: The dataset covers a wide range of temporal correlation and spatial interaction. From these patterns, researchers can build realistic occupancy behavior model. A preliminary analysis of the dataset can be found in [3].

4.1 Case study: HVAC control using occupancy data

In this case study, we investigate how we can utilize COD to design a setback schedule and control the power of the heater for the heating in a conference room. A model predictive controller (MPC) for the heater, implemented in OpenBuild[1], will take the setback schedule and dynamically adjust the power input. The results of our simulations are shown in Figure 3. The top figure is the occupancy profile of the room. The figure in the middle shows the temperature setpoint constructed from the occupancy trace (when unoccupied, the setpoint is reduced to 17°). The bottom figure shows the power input required when using a constant setpoint (which was the case for the actual conference) and the MPC solution.

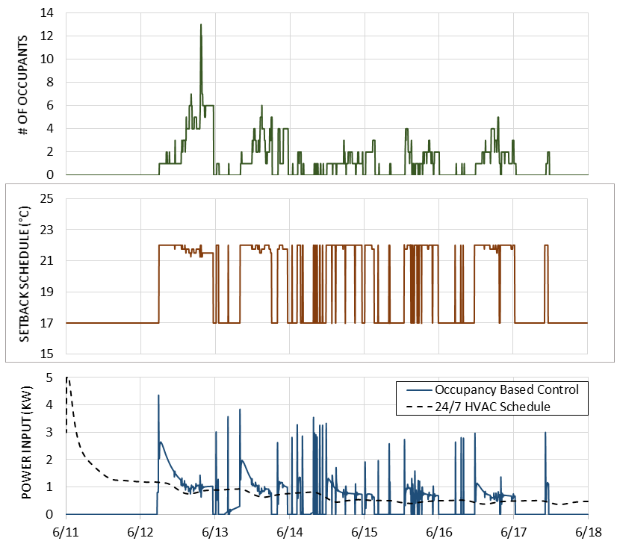

Figure 3: HVAC control for a conference room using occupancy information

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Department of Energy grant DE-EE0007682.

REFERENCES

[1] Yuvraj Agarwal, Bharathan Balaji, Rajesh Gupta, Jacob Lyles, Michael Wei, and Thomas Weng. 2010. Occupancy-driven energy management for smart building automation. In *Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building*. ACM, 1–6.

[2] Drury B Crawley, Linda K Lawrie, Frederick C Winkelmann, Walter F Buhl, Y Joe Huang, Curtis O Pedersen, Richard K Strand, Richard J Liesen, Daniel E Fisher, Michael J Witte, et al. 2001. EnergyPlus: creating a new-generation building energy simulation program. *Energy and buildings* 33, 4 (2001), 319–331.

[3] Kin Sum Liu, Sirajum Munir, Jonathan Francis, Charles Shelton, and Shan Lin. 2017. Long term occupancy estimation in a commercial space: an empirical study. In *IPSN*. 307–308.

[4] Kin Sum Liu, Elvin Vindel Pinto, Sirajum Munir, Jonathan Francis, Charles Shelton, Mario Berges, and Shan Lin. 2017. COD: A Dataset of Commercial Building Occupancy Traces. (Sept. 2017). https://doi.org/10.5281/zenodo.996587

[5] Sirajum Munir, Ripudaman Singh Arora, Craig Hesling, Juncheng Li, Jonathan Francis, Charles Shelton, Christopher Martin, Anthony Rowe, and Mario Berges. 2017. Real-Time Fine Grained Occupancy Estimation using Depth Sensors on ARM Embedded Platforms. In *Real-Time and Embedded Technology and Applications Symposium (RTAS), 2017 IEEE*. IEEE, 295–306.

[6] Kaiyu Sun, Da Yan, Tianzhen Hong, and Siyue Guo. 2014. Stochastic modeling of overtime occupancy and its application in building energy simulation and calibration. *Building and Environment* 79 (2014), 1–12.

[1] http://la.epfl.ch/openBuild