Room-level Occupant Counts, Airflow and CO_2 Data from an Office Building

Krzysztof Arendt
University of Southern Denmark
krza@mmmi.sdu.dk

Aslak Johansen
University of Southern Denmark
asjo@mmmi.sdu.dk

Bo Nørregaard Jørgensen
University of Southern Denmark
bnj@mmmi.sdu.dk

Mikkel Baun Kjærgaard
University of Southern Denmark
mbkj@mmmi.sdu.dk

Claudio Giovanni Mattera
University of Southern Denmark
cgim@mmmi.sdu.dk

Fisayo Caleb Sangogboye
University of Southern Denmark
fsan@mmmi.sdu.dk

Jens Hjort Schwee
University of Southern Denmark
jehs@mmmi.sdu.dk

Christian T. Veje
University of Southern Denmark
veje@mmmi.sdu.dk

ABSTRACT

The area of occupant sensing is lacking public datasets to baseline and foster data-driven research. This abstract describes a dataset covering room-level occupant counts, in-room ventilation airflow and CO_2 data from an office building. This dataset can among others be used for developing and evaluating data-driven algorithms for occupant sensing and building analytics.

CCS CONCEPTS

• **Information systems → Data cleaning**;

KEYWORDS

Data, Occupant Sensing, CO2

ACM Reference Format:

Krzysztof Arendt, Aslak Johansen, Bo Nørregaard Jørgensen, Mikkel Baun Kjærgaard, Claudio Giovanni Mattera, Fisayo Caleb Sangogboye, Jens Hjort Schwee, and Christian T. Veje. 2018. Room-level Occupant Counts, Airflow and CO_2 Data from an Office Building. In *The 1st Workshop on Data Acquisition To Analysis (DATA'18), November 4, 2018, Shenzhen, China.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3277868.3277875

1 INTRODUCTION

Sensing accurately the number of occupants in the rooms of a building has many applications including smart spaces, safety and evacuation, facility management and building operation. In terms of building operation accurate people counts can enable applications, such as, adaptive ventilation in rooms, occupant-based energy benchmarking, and model-predictive control of room setpoints. In all these applications the more accurately the number of occupants can be sensed the safer or more energy-efficient a building can become [1, 3]. This has fostered research around these topics based on using different types of sensor modalities and data sources. Camera sensors, ventilation airflow measurements and CO_2 sensors are some of the popular sensor modalities for such research [7]. However, a wide set of public datasets are missing that enable baseline studies of algorithms with long term ground truth information for different types of buildings. To address this we released a dataset covering occupant counts based on 3D stereo vision camera, damper openness to represent airflow into each room and CO_2 concentration measured by in-room sensors.

The data is obtained from a large teaching and office building at the University of Southern Denmark designed to be a living lab for data-driven research on buildings [5]. The building is 8,500 m^2, records an average of 1,000 occupants on normal weekdays and facilitates several types of staff and student activities. Room-based data was collected for four rooms where two are regular teaching rooms and two are study zones. The study zones have a mixed use for student activities, such as, project work and solving exercises. The teaching rooms are most of the time booked for lectures spanning two to four hours for each class. Therefore, the data includes a unique combination of important sensor modalities for room-level occupant sensing collected in both rooms with scheduled activities and rooms with unscheduled activities.

The occupant counts are collected by eight highly accurate PC2 3D stereo vision cameras from the company Xovis mounted over the two room entrances to each of the four rooms. An implementation of the PLCount algorithm [6] was used to clean and fuse the raw count data to room counts. Previous results with this type of sensor and the PLCount algorithm has demonstrated an accuracy of 0.075 RMSE compared to a manual ground truth [6]. Therefore, the accuracy of the occupant count data is several times more accurate than that of airflow and CO_2 based algorithms [7] and can therefore be used as ground truth. In-room airflow – measured as the damper openness – and CO_2 concentration measurements are available for the four rooms. Data are collected from the sensors via an external API of the building management system. The released CO_2 sensor data has not been cleaned and users should take care to address known issues with CO_2 data including offsets and drifts [7].

The released dataset contains data from fifteen days in the Spring of 2017. This dataset was selected from a longer data sample to

DATA'18, November 4, 2018, Shenzhen, China

ACM ISBN 978-1-4503-6049-4/18/11…$15.00
https://doi.org/10.1145/3277868.3277875

be free of data omissions or other collection errors. The temporal resolution of the data is minute-wise. The occupant counts are the estimated number of people present in the room, the duct-airflow is represented as the damper position openness in percentage of fully opened and CO_2 as the concentration measurements of the sensor in ppm. In total the data for each room-level sensor contains 21,600 readings which in total sums up to 259,200 readings for the four rooms and three data types. Due to the living lab study design the number of unique occupants monitored is not available. In total occupants spent 5,713 hours in the monitored rooms during the data collection. The released data is plotted in Figure 1 for each data type and room by overlaying daily profiles. The dark line represent the average of all the daily profiles per minute. The dataset also contains metadata about the four rooms including size and seating capacity and a Brick representation of the sensor instrumentation [2]. The most sensitive part of this data is that the counts from the lecture rooms represent class attendance for specific classes of the faculty. Therefore, the room identifiers and the actual days have been removed. A random day identifier has replaced the actual day and the data sample for each day is included in random order in the released dataset, but still using the same order for all the sensors. Additional labels have been added for the day of the week and holidays. The applied anonymization limits the use of the dataset to within the same day estimation and prediction tasks. However, the majority of published studies only considers the same day making the dataset usable for most studies.

The potential use of the data includes evaluation of models and algorithms for room-level occupant presence counting, prediction of occupant presence and models of occupant actions in rooms with scheduled and unscheduled activities [3, 6, 7]. The data might also be used for testing anonymization algorithms [4].

ACKNOWLEDGMENTS

This work is supported by the Innovation Fund Denmark for the project COORDICY (4106-00003B).

REFERENCES

[1] M. Angermann, M. Khider, and P. Robertson. 2008. Towards operational systems for continuous navigation of rescue teams. In *2008 IEEE/ION Position, Location and Navigation Symposium*. 153–158.

[2] Bharathan Balaji, Arka Bhattacharya, Gabe Fierro, Jingkun Gao, Joshua Gluck, Dezhi Hong, Aslak Johansen, Jason Koh, Joern Ploennigs, Yuvraj Agarwal, Mario Berges, David Culler, Rajesh Gupta, Mikkel Baun Kjærgaard, Mani Srivastava, and Kamin Whitehouse. 2018. Brick : Metadata Schema for Portable Smart Building Applications. *Applied Energy* (2018).

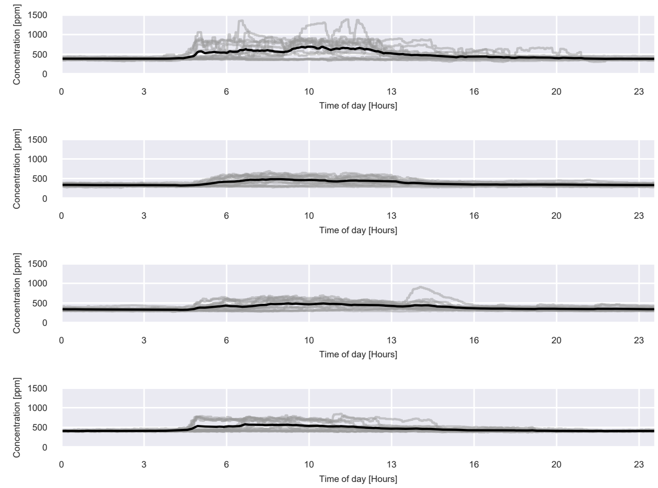
[3] Varick Erickson, Miguel A. Carreira-Perpinan, and Alberto E. Cerpa. 2011. OBSERVE: Occupancy-based system for efficient reduction of HVAC energy. In *IPSN 2011*. 258–269.

[4] Ruoxi Jia, Fisayo Caleb Sangogboye, Tianzhen Hong, Costas J. Spanos, and Mikkel Baun Kjærgaard. 2017. PAD: protecting anonymity in publishing building related datasets. In *BuildSys*.

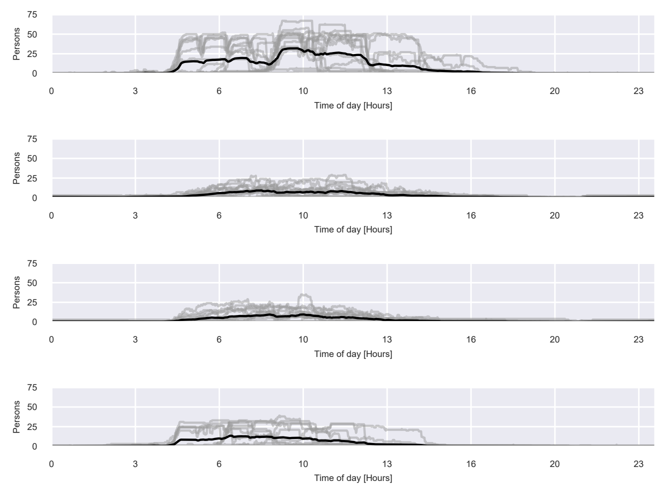
[5] Muhyiddine Jradi, Fisayo Caleb Sangogboye, Claudio Giovanni Mattera, Mikkel Baun Kjærgaard, Christian Veje, and Bo Nørregaard Jørgensen. 2017. A World Class Energy Efficient University Building by Danish 2020 Standards. *Energy Procedia* 132 (2017), 21 – 26.

[6] Fisayo Caleb Sangogboye and Mikkel Baun Kjærgaard. 2016. PLCount: A Probabilistic Fusion Algorithm for Accurately Estimating Occupancy from 3D Camera Counts. In *ACM BuildSys 2016*.

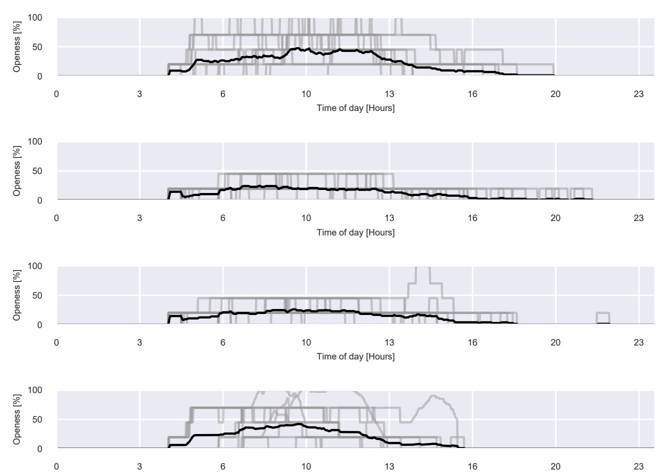
[7] Fisayo Caleb Sangogboye, Krzysztof Arendt, Ashok Singh, Christian T. Veje, Mikkel Baun Kjærgaard, and Bo Nørregaard Jørgensen. 2017. Performance comparison of occupancy count estimation and prediction with common versus dedicated sensors for building model predictive control. *Building Simulation* 10, 6 (01 Dec 2017), 829–843.



(a) CO_2 concentration



(b) Occupant counts



(c) Damper position

Figure 1: Daily profiles of the released data