MODES: Multi-sensor Occupancy Data-driven Estimation System for Smart Buildings

Hamid Rajabi, Zhizhang Hu, Xianzhong Ding, Shijia Pan, Wan Du, and Alberto Cerpa*
{hrajabi2,zhu42,xding5,span24,wdu3,acerpa}@ucmerced.edu
University of California, Merced
Merced, California, USA

ABSTRACT

Buildings account for more than 40% of the energy US primary energy consumption. Of all the building services, heating, ventilation, and air-conditioning (HVAC) account for almost 50% of that energy use. Despite all the resources used, many users are not satisfied with the comfort conditions in buildings. The main problems for this lack of balance between energy use and quality of comfort are the lack of occupancy information, real user comfort feedback, and easily built zone thermodynamic models available to the Building Management Systems (BMS). In our work, we focus on occupancy sensing. While occupancy sensing is very important and there are multiple different sensing technologies used to address this issue, a precise and reliable measurement of occupancy remains elusive.

In this paper, we propose MODES, a Multi-sensor Occupancy Data-driven Estimation System for Smart Buildings. Leveraging on two different state-of-the-art sensing techniques available in the literature (vibration and thermal sensors), both being capable of counting the number of occupants in any particular zone. The two occupancy estimations are then fused using a data-driven optimization process for sensor fusion to create an improved estimate. This newly updated estimate is further used together with a data-driven occupancy model as input of a particle filter to provide an even more accurate estimate. We tested the system in a commercial building under realistic conditions using real experimental occupancy data traces with users doing their daily routines. We showed that MODES can improve occupancy estimation by 40% from vibration sensors, 19% from thermal sensors, and 30% from state-of-the-art sensor fusion schemes. Moreover, we show that this is possible with minimum data training requirements, needing 7 days of training data to train the fusion system. We also run several EnergyPlus simulations using an occupancy-driven HVAC controller under different occupancy errors to show the impact that more accurate occupancy sensing schemes can have on the overall energy usage and quality of comfort and air ventilation. Our study shows that MODES can save up to 77% of energy use in a building while improving the quality of comfort by 10%. [1]

[1]The full dataset of this study could be found on: https://doi.org/10.15783/08kw-7n74

e-Energy '22, June 28-July 1, 2022, Virtual Event, USA

ACM ISBN 978-1-4503-9397-3/22/06.
https://doi.org/10.1145/3538637.3538852

CCS CONCEPTS

• **Computer systems organization → Sensor networks**; **Embedded systems**; *Redundancy*; Robotics; • **Computing methodologies → Supervised learning by classification**.

KEYWORDS

occupancy estimation, HVAC systems, thermal occupancy sensor, vibration occupancy sensor, late sensor fusion, particle filter

ACM Reference Format:
Hamid Rajabi, Zhizhang Hu, Xianzhong Ding, Shijia Pan, Wan Du, and Alberto Cerpa. 2022. MODES: Multi-sensor Occupancy Data-driven Estimation System for Smart Buildings. In *The Thirteenth ACM International Conference on Future Energy Systems (e-Energy '22), June 28-July 1, 2022, Virtual Event, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3538637.3538852

1 INTRODUCTION

People spend 87% of their time inside buildings [15, 34], making buildings an important part of our lives. To be comfortable, we spend a significant amount of energy. In the US alone, buildings account for 40% of energy usage [2, 16], and of that 50% of energy goes to heating, ventilation, and air conditioning (HVAC) [31, 35]. Despite this expense, more than 75% of the occupants in commercial buildings are not satisfied with their thermal comfort [23]. This is mostly due to a lack of information, more specifically, where the occupants are and will be (occupancy), what they prefer (comfort), and how to condition the different zones (thermodynamic models) [50]. Concerning occupancy, currently, buildings' zones are conditioned based on time constraint assumptions (e.g. occupied and unoccupied hours), whether the zones are occupied or not, wasting significant amounts of energy. Building Management Systems (BMS) do not have information on *where* the occupants are and will be in the building [50]. Having detailed occupancy information about the number of people in real-time and in the near future for each zone allows to control the HVAC systems in a much more efficient manner [4, 22, 24], conditioning zones for temperature only when occupied, and adjusting ventilation rates based on the actual number of occupants. Such HVAC control systems have been investigated in [6] to lower the peak of energy demand through adaptive HVAC schedules for each zone. To tackle this issue, multiple different sensing modalities have been used to measure occupancy, including CO_2, PIR, ultrasonic, image, sound, EM signals, power meters, and computer applications [36, 52]. With a plethora of multiple sensing modalities available, sensor fusion techniques [17, 19] that use multiple sensor inputs to try to obtain better occupancy estimation seem like a promising path. The majority of related work in the multi-sensor occupancy estimation concentrates on *early sensor fusion and classification*, where different data features are extracted from the raw sensor values to produce a combined implicit occupancy estimate [10, 38]. The inaccurate raw data measured by some commonly available environmental sensors (e.g., electrical load meters, CO_2, temperature, humidity sensors, etc.) is then modeled and combined to achieve accurate occupancy estimation. This type of early fusion, however, may not be computationally-effective, as it may be difficult to handle when the number of environmental features increases [45].

In this paper, we take a different approach. Instead of trying to combine the raw signals of all the different sensor modalities for occupancy sensing, we use the processed output of each occupancy sensor to produce a more accurate occupancy sensing estimation. Using a *late sensor fusion and classification* scheme, allows us to use the occupancy sensors *as is*, with no modifications, and use their processed occupancy estimation in our sensor fusion pipeline to obtain a more accurate occupancy estimation than any of them individually. The idea is to leverage the synergistic approach of multiple different sensors' error modes to complement each other, with the final result being better than the sum of their parts.

We developed MODES, a Multi-sensor Occupancy Data-driven Estimation System for Smart Buildings. The system was tested with two different state-of-the-art occupancy sensing technologies available in the literature, using vibration sensors [42] and thermal/PIR sensors [7]. MODES' collection module consists of thermal and vibration sensors deployed on the building zones' ceilings and floors, respectively. Both sensors provide time-series occupancy estimations using some classification algorithms applied to the raw sensor data. MODES combines the two occupancy streams through a Data-driven Optimization-based Weighted Average (DOWA) algorithm, to calculate the optimal fusion weights between the two data streams. This occupancy estimation is further refined by using a non-linear particle filter, capable of tracking multiple concurrent hypotheses, which uses an occupancy model based on the Blended Markov Chain (BMC) technique [21]. Our initial results show that MODES' occupancy estimation is the best using multiple metrics.

The main contributions of this work are as follows:

- We developed MODES, a late-fusion technique that uses occupancy estimation from different occupancy sensing systems (using different sensing modalities), and combines them to obtain a more accurate occupancy estimation result.
- We deployed two state-of-the-art occupancy sensing systems based on vibration sensors [42] and thermal sensors [7], on an office building, testing their performance over two weeks.
- We analyzed the system performance, showing the breakdown of each component in the processing pipeline (vibration, thermal, DOWA, and particle filter), as well as a state-of-the-art late fusion technique [13] to highlight the benefits of MODES under two different scenarios: low zone occupancy (up to two occupants), and high zone occupancy (up to eight occupants).
- We performed energy and comfort simulation analysis using EnergyPlus [1] and an occupancy-based HVAC controller to measure the impact of our improved occupancy estimation, highlighting its impact on the energy savings and quality of comfort trade-offs for buildings operations.

2 RELATED WORK

2.1 Background

User-based sensing systems, where the users carry tags or devices have been traditionally used to measure room-level occupancy in the past [11, 12, 26]. Although such systems may have acceptable accuracy and performance, they may not be always practical as people always need to carry a device with them, and therefore they are intrusive. In addition, they suffer in performance when the user forgets to carry the tag/device.

User-free schemes [5, 7, 40, 55] do not suffer from the problems mentioned above. There are many different sensing technologies, including video-based sensing systems [8, 33, 48], CO_2 sensors [32, 41, 55], vibration systems [39, 42, 43], WiFi-based systems [37, 54], and thermal-based occupancy sensing modalities [7, 40, 53]. Video-based systems have limitations in Non-Line-Of-Sight (NLOS) environments, in the dark, through smoke, or walls. They are also computationally intensive to account for complicated image processing and have deployment costs. In addition, due to privacy concerns, they cannot be deployed inside the zones where people do their daily tasks and their deployment is limited to public hallway spaces inside the buildings. To control ventilation, CO_2 sensors are commonly deployed in buildings. However, they suffer from calibration issues, and the delay between the arrival and departure of people and the associated changes in CO_2 level buildups, which make their use difficult for real-time building control. Moreover, they are very sensitive to the deployment positions within the zone (height and specific deployment location) [27]. Many WiFi-based occupancy estimation systems need users' movement to be able to locate them [5]. Therefore, they are less efficient when users are in more static environments. They also suffer from poor coverage areas, especially in rooms with high-density furniture or machines, firewalls, and electromagnetic interference from other devices.

Groß et al. in [28] presented a Pugh Matrix providing a visual comparison between individual state-of-the-art occupancy sensing techniques based on a set of features and requirements. Each feature in the matrix is assigned to pick a rating of -1, 0, or 1, in the order of the worst-to-best state. Those rating values were then summed up to deliver an evaluation metric for each occupancy estimation technique. The features evaluated were accuracy, detection range, sensor coverage area, delay, computation costs, user's privacy concerns, ease of installation, and costs. Appendix A shows a summary of their results. The most two promising techniques are the Heat-map (i.e., thermal-based) and Vibration sensing systems. Therefore, we selected these two suitable state-of-the-art solutions for our evaluation. We picked them since they have complementary strengths. We provide more details about them below.

2.2 Vibration-based Occupancy Sensing

Various infrastructure vibration-based sensing systems have been developed to capture indoor occupant information. Their main mode of operation is to try to capture the interactions of people with the building. When occupants use the building for their daily tasks (e.g., walk, move, sit, etc.) their actions induce the surrounding objects (e.g., floor, table, wall) to vibrate, and this vibration information can be captured. This vibration contains the information of both the structure of the vibrating objects and the source

Figure 1: ThermoSense node deployed in the ceiling. The enclosure includes a thermal array and a PIR sensor.

Figure 2: BOES vibration occupancy sensing system, deployed on the floor to capture footsteps.

of the vibration, i.e. the occupant action. As a result, the captured vibration can be used to infer the building occupant information, such as presence [42], occupancy [43], identity [44], activity [30], and location [39]. The advantages of this sensing modality include NLOS sensing without much privacy concerns. The main drawbacks are its limited capability of providing an accurate occupancy count for a large group of occupants, especially in high occupancy environments (e.g. more than 4) in a particular zone.

2.3 Thermal-based Occupancy Sensing

The main idea behind thermal-based sensing is the temperature differential between occupants and their surrounding environment, to infer the presence of occupants and even the number of them [7, 40, 49, 53]. These systems are capable of detecting occupants in large groups both in static (e.g. sitting) and dynamic (e.g. moving) positions. They do require LOS for detection, but differently from camera-based systems, they are significantly less intrusive, since they just measure temperature instead of a full video. This means that they can be deployed inside the zones where people work. In our work, we utilize the ThermoSense occupancy sensing system [7]. It is a combination of a PIR sensor for binary detection of occupancy, together with a temperature array that allows counting the number of people in a zone.

2.4 Sensor Fusion for Occupancy Estimation

When multiple occupancy sensors are available, it is possible to combine their values for a more accurate occupancy sensing estimation using sensor fusion techniques. In the literature, sensor fusion is done by early fusion or late fusion.

Early fusion methods are based on supervised learning techniques, with feature vectors extracted from the raw signals of all the sensors to classify the desired output [46]. The major weakness of these approaches is the high complexity due to the high number of features, requiring significant feature engineering work. In addition, the classifier, and features must be trained for the very specific sensors used, meaning the need to develop and train a new early fusion method when a new sensing modality is added [9, 18, 20, 29].

Late fusion methods have the advantage of not requiring changes to the occupancy sensors processing pipelines, using only the occupancy estimation values that each sensor system provides. This means that the number of features processed by each sensor is reduced, reducing complexity. The disadvantage is that features from multiple sensors cannot be globally optimized since we are working with the processed occupancy output of each sensor. The Occupancy Aware Clustering (OAC) work [13] is an example of a late fusion technique, during which the occupancy estimations based on different sensors are fused with a linear regression classifier as part of its late fusion phase. In our work, we use a combination of data-driven occupancy weighted average (DOWA) together with a non-linear particle filter that allows us to get a more accurate estimate. Late fusion can correct for the shortcomings of each commercial sensor input when there is no access to the raw data.

3 MODES SYSTEM OVERVIEW

Figure 3 shows the overall MODES processing pipeline. The inputs to the left of the figure are the vibration [42] and thermal sensors [7], which are described in more detail in the sections below. Each of these sensors performs an independent data-driven classification based on the raw signals sensed by the sensors, and each sensor outputs a discrete occupancy estimation value. This data is used as input to the Data-driven Optimization-based Weighted Average (DOWA) which fuses both occupancy estimates. DOWA's output is further used as input of a particle filter, which uses a data-driven occupancy model (BMC) to further refine the occupancy estimation. The final output of the process is a discrete value with a more accurate occupancy estimation in the zone.

3.1 ThermoSense

We use the ThermoSense [7] thermal occupancy sensor system as one of our sensor modalities. Figure 1 shows a picture of the system deployed on the ceiling. This system has 2 sensors, a PIR sensor, and a thermal array sensor. The PIR helps detect the movement of occupants and triggers the more energy-expensive thermal array, which can count the number of occupants underneath. It consists of an 8 by 8 thermal sensor that can cover an area of approximately 7 square meters (depending on the ceiling height). The sensors perform an internal complex processing pipeline, including background subtraction to determine the active pixels in the thermal array (those that might detect a human), connected components to determine the different blobs, and then it uses a classifier (e.g., linear regression, KNN or ANN) to determine the total number of occupants based on certain features, such as the total number of active thermal pixels, the total number of blobs, the size of the largest blob and more. The final discrete output is the total number of occupants detected by the sensor. Multiple sensors are deployed in each zone depending on the size of the area to be covered.

3.2 BOES

We use the Building Occupancy Estimation System (BOES) [42] as our vibration-based occupancy sensor in our work. Figure 2 shows a picture of the system deployed on the floor. BOES is a non-intrusive occupancy monitoring solution that uses sparsely-deployed vibration sensors. The system captures vibration signals on the floor, then detects footstep events, and finally, recognizes significant events based on observed walking traits. The covering area is measured to be between 2 and 4 feet radius from the sensor. The system has three major modules: event detection, localization, and tracking. The vibration sensors forward the floor vibration

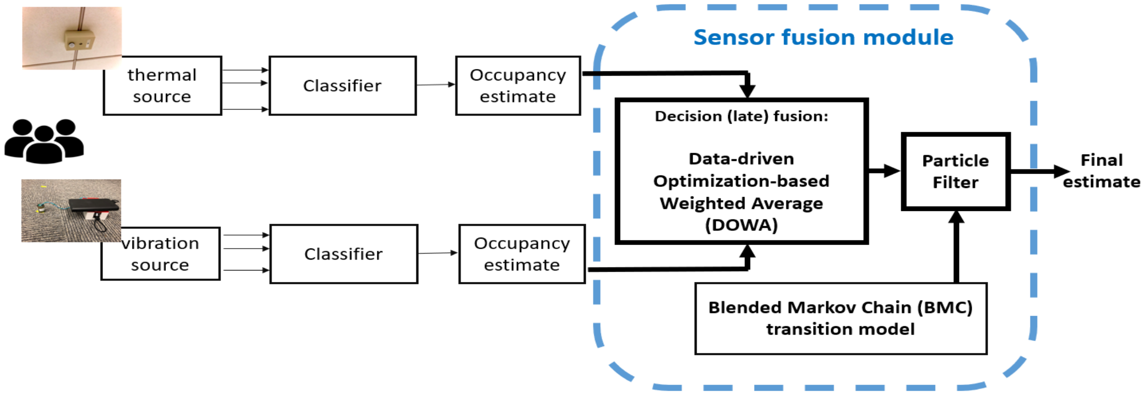

Figure 3: MODES Processing Pipeline.

velocity to the step event detection module. Whenever a step event is detected, it is forwarded to the localization unit. This unit locates the sequence of step events by comparing the time of peak step event energy at each sensor. Finally, the tracking unit seeks for velocity change events between two or three consecutive step events. Then, the tracking unit combines the velocity events with the localized step events obtained by the localization unit to update the final occupancy count. The final discrete output is the estimated occupancy count in the zone.

3.3 Data-driven Optimization-based Weighted Average (DOWA)

The sensor fusion process in MODES starts with a weighted average between two sensor inputs occupancy estimates. Each weight is defined as a quadratic function of the associated input data. Obtaining the optimal weight is a supervised learning problem, with the optimal coefficients for each weight function obtained through an optimization process that minimizes the least square error to ground truth data. The data requirements for convergence to good parameter estimation are explored in the following section. We called it a Data-driven Optimization-based Weighted Average or DOWA. Table 1 describes each parameter in the algorithm. The process to obtain the optimal parameters is as follows.

The estimated combined sensor data, $\hat{z}_i$, is obtained through a weighted average as follows,

$$\hat{z}_i = w_t z_{t,i} + w_v z_{v,i} \tag{1}$$

where the weight functions, w_t and w_v, are defined as follows,

$$\begin{aligned} w_t &= a z_{t,i}^2 + b z_{t,i} + c \\ w_v &= a' z_{v,i}^2 + b' z_{v,i} + c' \end{aligned} \tag{2}$$

Note that both weight functions are considered to output positive reciprocal values (i.e., they sum to one). They are obtained through optimizing a non-linear Least Square Regression as the objective function,

$$\underset{a,b,c,a',b',c'}{\arg\min} \sum_{i=1}^{T} |\hat{z}_i - z_i|^2 \tag{3}$$

Table 1: Table of notations for DOWA algorithm.

parameter	description
$z_{t,i}$	occupancy estimate by thermal sensor at time i
$z_{v,i}$	occupancy estimate by vibration sensor at time i
z_i	ground truth occupancy at time i
$\hat{z}_i$	estimated combined sensor data at time i
w_t	normalized weight function for thermal input
w_v	normalized weight function for vibration input
a, b, c	optimal weight coefficients for thermal input
a', b', c'	optimal weight coefficients for vibration input

where T is the total number of data points. The linear constraints to the objective function are listed below:

$$\begin{aligned} (i) &\quad w_t + w_v = 1, \\ (ii) &\quad 0 \leq w_t \leq 1, \\ (iii) &\quad 0 \leq w_v \leq 1. \end{aligned} \tag{4}$$

Note that the values of $z_{t,i}$ and $z_{v,i}$ may be the same in some time samples. In such cases, we do not need an optimization process to estimate the $\hat{z}_i$, as it can be equal to either one of the sensor inputs. However, optimization is needed for the rest of the samples i.e., $z_{t,i} \neq z_{v,i}$. By defining non-linear weights, we can capture more complex relationships of the different accuracy that the sensors may have depending on the total number of occupants being estimated.

Through this methodology, we can obtain an improved occupancy estimate where the occupancy estimate would be closer to the sensor data which is more similar to the ground truth. However, not all errors can be corrected, since *both* occupancy sensors may produce incorrect occupancy estimates. While DOWA can minimize the error by weighing more heavily on the sensor that is more accurate for that particular occupancy output, we cannot correct all the occupancy errors.

3.4 Particle Filter

To further improve the occupancy estimation accuracy, a particle filter is used. The samples of the posterior occupancy are so-called *particles* and are denoted as $X_t = x_t^1, x_t^2, ..., x_t^m$. The particles are

Algorithm 1: Particle Filter (X_{t-1},z_t)

```
initialization: X̄_t = X_t = 0;
for m = 1 to M do
    sample x̄_t^m from Occupancy Model (x̄_{t-1}^m, t);
    add x̄_t^m to X̄_t ;
    get β_t^m from Measurement (z_t, x̄_t^m);
end
β̄_t^m = Normalize (β_t^m), m = 1, ..., M;
for m = 1 to M do
    draw x_t^m with probability β̄_t^m from X̄_t;
    add x_t^m to X_t ;
end
return X_t
```

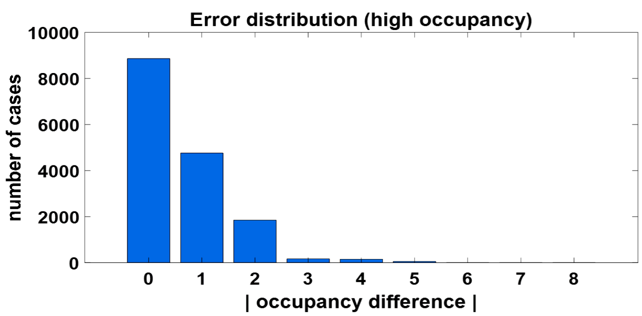

Figure 4: Error distribution of the input measurement.

used to estimate the distribution of the posterior occupancy state. The process is shown in Algorithm 1. M denotes the number of particles in the set X_t, and time t represents the time in seconds. The algorithm includes three major phases described below.

3.4.1 ***Phase 1: sampling from the occupancy model.*** In our work, the occupancy model is represented by a Blended Markov Chain (BMC) as described in [21]. During this phase, M samples are drawn from the BMC model. Each one of the states x_{t-1}^m can jump to a possible successor state with probability $p(\hat{x}_t^m|x_{t-1}^m)$; $m = 1, 2, ..., M$.

3.4.2 ***Phase 2: calculation of the particle weights.*** The weight β_t^m is obtained from the measurement model. This model is calculated from the distribution of the difference between the ground truth and the occupancy estimate from the sensor (i.e., DOWA's output) in the training set (Figure 4). The weight β_t^m of a particle is obtained by calculating the absolute difference between the model estimation $\bar{x}_t^m$ and the DOWA's output z_t. Finally, the weight $\bar{\beta}_t^m$ in a particle m is obtained by dividing the number of cases of that absolute difference by the total number of cases.

3.4.3 ***Phase 3: re-sampling.*** We draw with replacement M samples from the particle set $\bar{X}_t$, which is proportional to the weight β_t^m. Then, this new particle set X_t, is the desired one to determine the final occupancy of the room at each time step.

The final output of the particle filter is the most accurate occupancy estimate corrected by the best optimal occupancy estimation values using DOWA and further corrected by using a data-driven occupancy-based model.

4 PERFORMANCE EVALUATION

In this section, we proceed to experimentally evaluate the performance of our system, including the performance breakdown of all the MODES components, i.e. vibration, thermal, DOWA and particle filter, as well as the Occupancy Aware Clustering (OAC) [13], which is a state-of-the-art late fusion technique.

4.1 Experimental Setup

We deployed 9 thermal and 4 vibration sensor motes in a university research lab. The lab has an area of approximately 51 m^2 with a rectangular shape, containing 8 cubicles total, 4 on each side of the rectangle's long sides, a meeting table in the middle with 6 chairs around it. There are also file cabinets near the entrance of the lab. During the experimental days, the room started being occupied at around 10:00 am, and the last occupants left at around 5:00 pm. We categorized our experimental data into two scenarios: one of low occupancy (up to 2 occupants) and another one of high occupancy (up to 8 occupants). We recorded 14 days of weekday data (almost 3 weeks), having 11 days of low and 3 days of high occupancy.

The thermal sensors were deployed on the ceiling covering the entire lab. To avoid missed detection due to blind spots, we deployed the thermal sensor coverage areas very close to each other. Therefore, sensor overlapping is inevitable when an occupant is located at the edge of two adjacent coverage areas. In this case, it may be counted twice. The vibration sensors were deployed at the entrance of the lab, acting as a vibration turnstile for all the occupants coming in/out of the lab. The thermal sensors can process the data locally and transmit with a low-power 802.15.4 radio an occupancy count, which is then aggregated in our back-end server. The vibration sensors do not have wireless transceivers and they were connected to the network with Ethernet cables. We also run a time synchronization protocol to loosely synchronized all the sensors to less than 150 ms, so we can temporally correlate the events of the different sensors. We were trying to estimate occupancy every second, so we got around 400,000 data points in total. Note that this was done to get statistically significant results. For HVAC building control, the occupancy estimate should be ~5 to 15 minutes, to match the actuation interval commonly used in buildings.

4.2 Evaluation Metrics

We evaluated our system using the following performance metrics.

4.2.1 *Accuracy.* The occupancy estimates by each scheme (i.e. MODES component) were classified into discrete values. The basic accuracy metric, in this case, would be the number of correctly estimated occupancy samples over the total number of examined data points over time. In terms of positives and negatives, it can also be defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

where TP, TN, FP, and FN are respectively denoting true positive, true negative, false positive, and false negative predicted samples.

4.2.2 *F-score.* Since we have a multi-class classification problem in this work, we evaluated the estimates based on the F-score (F-measure) [47]. F-score is used for binary classification, however,

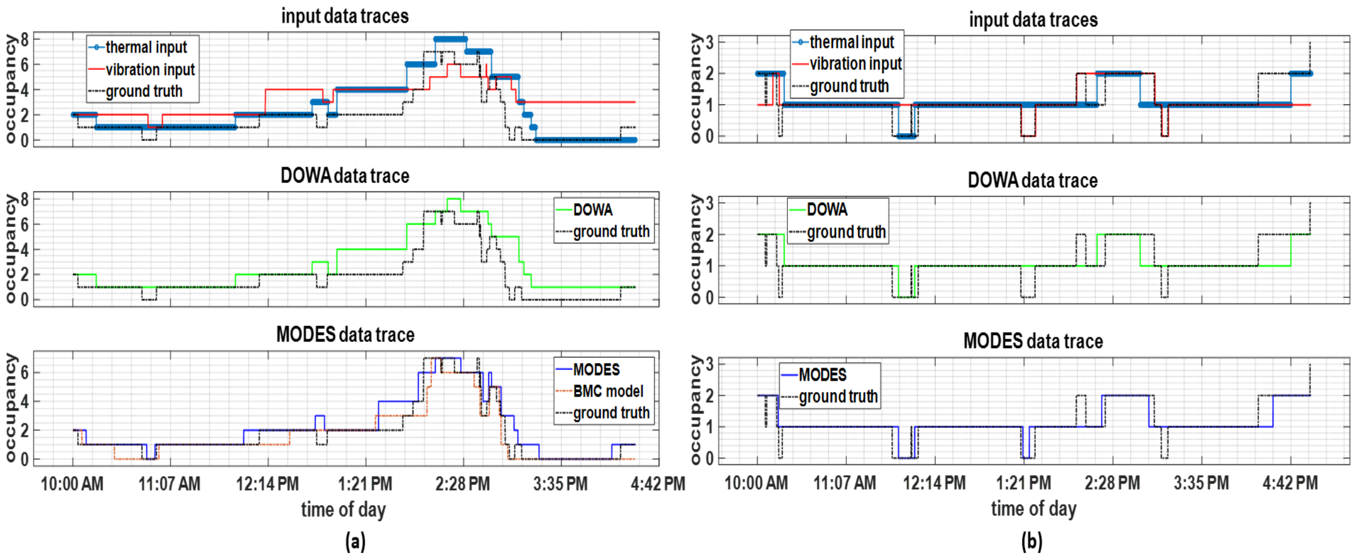

Figure 5: Data traces showing the sensor inputs and processed data for one day (a) high occupancy and (b) low occupancy.

we can treat our classifier as a One-vs-rest to test its accuracy. The F-score is defined as follows:

$$F - score = 2.\frac{precision * recall}{precision + recall} \quad (6)$$

where $precision = TP/(TP + FP)$ and $recall = TP/(TP + FN) = sensitivity$. The F-score is a real value between 0 and 1, and the closer to 1, indicates better precision and recall.

4.2.3 Over/Under-counting (OCR & UCR). We also investigated the over-counting and under-counting ratios. The over-counting ratio (OCR) is defined as the ratio of the number of cases in which the estimated occupancy is strictly larger than the corresponding true case, over the total number of examined cases. Similarly, the under-counting ratio (UCR) is the ratio of the number of cases in which the estimated occupancy is strictly smaller than the corresponding true case, over the total number of examined cases. Mathematically, the accuracy, OCR, and UCR ratios for a specific data scheme should sum to one, as the accuracy represents the correctly estimated samples and OCR/UCR represents the falsely estimated samples. The OCR/UCR metric is useful to determine if there is any bias in the occupancy estimator. It is also useful for the interpretation of the energy and quality of comfort analysis done later in the paper.

4.3 Exploratory Analysis

Figure 5 shows two examples of time-series occupancy in two days, for high (on the left) and low occupancy (on the right) scenarios. The top figures show the occupancy values for the different raw sensor values for both the thermal and vibration sensors, and compare them with the ground truth values. The middle figures show the DOWA corrected output, and finally, the bottom figures show the final MODES occupancy estimate.

DOWA combines the two input data streams coming from the thermal and vibration sources and then outputs discrete occupancy values. In these two cases, DOWA assigns the thermal and vibration weights of 0.7920 and 0.2080 for low occupancy, and 0.5414 and 0.4586 for the high occupancy case, with DOWA's output expected to be closer to the thermal source but with different coefficients (i.e. not constant over the input range). DOWA is a developed version of the traditional weighted sum since it prevents the average from floating outside of the two sources' occupancy ranges. That is observed in the range [~3:00 PM, ~end] of Figure 5-(a) where the vibration system is constantly over-counting by 3 occupants, and then, by adding the thermal scores, which are more precise, DOWA provides an occupancy score with smaller over-counting i.e., by only 1 occupant. Similarly, the thermal system experiences over-counting in the range [~2:10 PM, ~2:50 PM] of the high occupancy scenario, where the vibration system under-counts in that same range. That can be explained by the degraded performance of the vibration system in detecting a larger group of people (up to 7) and the thermal system overlapping in the sensing coverage (e.g. a person at the edge of two thermal sensing areas detected by both sensors). DOWA tries to smooth this score as well, however, the output has still some over-counting due to the larger weight assigned to the thermal input. The final estimate provided by the particle filter is more accurate, where the output of DOWA is further combined with the occupancy results obtained by the BMC model. As it is clear in both scenarios (specifically in high occupancy), the over- and under-counting sensor incidents have mostly been corrected by the particle filter. Hence, it is expected that the particle filter delivers the best accuracy among the four data schemes, which is the final accuracy for MODES. An example of how the BMC occupancy model may help the particle filter in improving occupancy accuracy is shown in Figure 5-(a) bottom figure. On the one hand, in the range [~ 11:50 AM, ~ 12:05 PM] where DOWA experiences an absolute occupancy difference of 1, the particle filter assigns more weight to the measurement model compared to the BMC model, since there is a large number of total cases in the error model with an absolute difference of 1 (see Figure 4). Therefore, it is expected for the particle filter to be more similar to the measurement model in this specific range. On the other hand, in the

Table 2: Performance evaluation for 5 different occupancy data schemes i.e. thermal and vibration occupancy sources, ordinary least squares (OLS), DOWA, and MODES, under the high and low occupancy scenarios (HO and LO).

		MODES	DOWA	OLS	thermal	vibration
Accuracy	HO	0.73	0.56	0.52	0.54	0.33
	LO	0.85	0.65	0.54	0.64	0.51
OCR	HO	0.21	0.40	0.32	0.41	0.46
	LO	0.10	0.19	0.03	0.19	0.33
UCR	HO	0.06	0.03	0.15	0.03	0.20
	LO	0.05	0.15	0.41	0.15	0.15

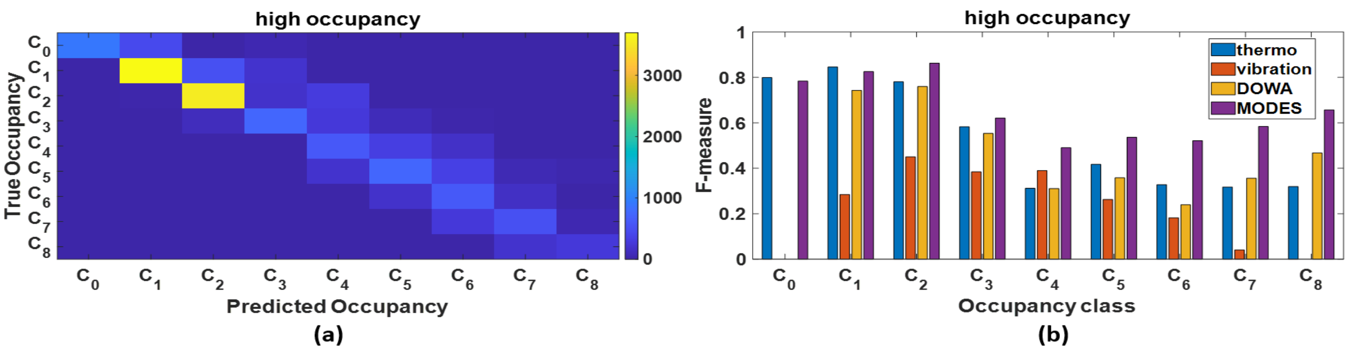

Figure 6: (a) MODES confusion matrix, and (b) MODES F-measure results.

short-range of [~2:40 PM, ~2:45 PM] where the error model experiences an absolute difference of 4 (in which there are few cases), the particle filter's output relies on the BMC and assigns more weights to it. Hence, the output is closer to the predicted by the BMC model, as is clear in the figure.

4.4 Accuracy

Table 2 presents the classification accuracy, OCR, and UCR as performance metrics evaluated for each data scheme. We also evaluated the Ordinary Least Squares (OLS) linear regression occupancy model examined in [13, 51] as a late fusion state-of-the-art technique for comparison. To address the multicollinearity issue in multiple linear regression models, we applied the Ridge Regression to make the OLS more robust against inaccurate estimates in regression coefficients [3]. Ridge regression applies a penalty on the size of coefficients and then minimizes a penalty residual sum of squares,

$$\arg\min_{\omega} ||Z_{t,v}\omega - Z||^2 + \alpha||\omega||^2 \tag{7}$$

where ω, $Z_{t,v}$, Z, and α denote regression coefficients, thermal and vibration occupancy vectors, ground truth data, and complexity parameter. The complexity parameter, α, is a positive value such that the greater value we assign to it, the more robustness against collinearity we achieve. In this analysis, α is selected to be 10 for both LO and HO scenarios through trial and error. In our case, the DOWA component obtains its optimal weight coefficients through a non-linear least squares regression, which is expected to be more flexible and fit better compared to the ridge regression. Furthermore, the optimization process in DOWA is constrained to deliver an occupancy count in the range of two input data streams (i.e., it does not float outside that range), while the traditional linear regression, as a late fusion technique, fails to consider that. Hence, this lack of range limit may cause additional inaccuracy for the OLS linear regression. The associated accuracy values in Table 2 reflect this issue. It is also clear that the final estimate of MODES is improved over DOWA. Considering a low accuracy of 0.33 for one of the sources (i.e. vibration), a final accuracy of 0.73 can be acceptable for a late fusion sensing module like MODES.

Based on Table 2, the occupancy errors by both thermal and vibration sources have experienced more over-counting in both high and low occupancy cases. The OLS has significant OCR and UCR values in both scenarios. Comparing these values with the ones from DOWA (which is closer to at least one of the input sources), demonstrates the effect of optimization constraints in DOWA that restricts the OCR/UCR values to being closer to one of the sources. With closer DOWA's results to the thermal sensor, we can estimate that DOWA has assigned a higher weight to the thermal sources with this dataset. Overall, MODES with our dataset experiences higher OCR values in both high and low occupancy scenarios compared to the UCR values.

Figure 6-(a) represents the corresponding confusion matrices of the MODES for the high occupancy scenario. Generally, the performance in lower occupancy classes e.g., C_1 and C_2, outperforms the one in high (or medium) levels. The reason for this behavior is in the pattern of each of the sources. Figure 6-(b) represents the entire behavior of multiple schemes in each occupancy class based on the F-score metric for high occupancy cases. Similar to the confusion matrix, the better performance of MODES is clear in the low occupancy classes. The very low value of DOWA in C_0, together with a high value for MODES is due to the BMC model in the particle filter, which is supposed to be assigned with a larger weight for this occupancy class.

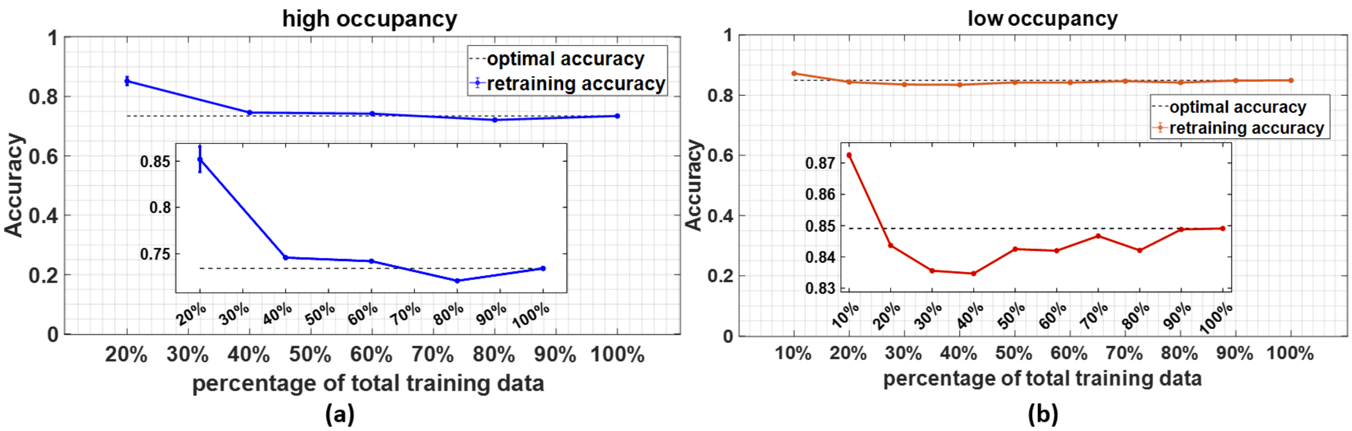

Figure 7: Accuracy as a function of the training data (with a zoom-in view) for (a) high occupancy, and (b) low occupancy.

4.5 Training Data Size

While the system accuracy is an important performance factor, the amount of training data required for a data-driven system is also important, since obtaining accurate ground truth data is a very expensive operation. Due to the non-linear nature of MODES (both in DOWA and particle filter components), it is expected to have high variance in the occupancy predictions, while at the same time being able to capture relevant relations between features and target outputs. In general, this comes at the expense of requiring more training data.

Figure 7 shows how the performance of MODES changes with a larger training set in case of high and low occupancy scenarios over the entire 3 and 11 days of experiments, respectively. We measure accuracy as a function of total training data, T, in each scenario. As Figure 7 shows, about 50% of total training data is required for the algorithm to get closed to converge to a stable accuracy, in both scenarios. The accuracy for the high occupancy scenario is lower than that of the low occupancy by ∼12%, and that proves the higher accuracy of MODES in the lower number of occupants case. In both occupancy scenarios, the initial accuracy is higher than the stable value. The reason for the high occupancy scenario is the initial data traces always start with fewer occupants in the day, with accuracy dropping as more occupants arrive at the lab. For the low occupancy scenario though, the initial accuracy is only a few percentage points from the stable accuracy value. Furthermore, it happens with a smaller data set (i.e., 10%) and converges faster for the high occupancy scenario. We note that for more than 20% of training data, the low occupancy accuracy has a small fluctuation of around ∼1% from the final stable accuracy. The very small standard deviation bars would also prove this fact. Accuracy stability is reached with about 40% percent of total data in high occupancy cases. In general, we see that the low occupancy scenario can reach stable accuracy levels with a lower size of T compared to the high occupancy case.

5 ENERGY ANALYSIS

In this section, we try to measure the energy and quality of comfort impact that more accurate occupancy estimation can have on HVAC building control using an occupancy-based controller. The energy analysis was simulated not only based on the time-series occupancy information for four data schemes but also based on the amount of over/under-heating/ventilation as one important input factor. We split the analysis down to over-counting and under-counting scenarios to study the effect of over/under-counting on energy use, temperature conditioning, and ventilation. The input occupancy to the EnergyPlus simulator is provided by the BMC occupancy model. The input model to the EnergyPlus includes 1589, 3496, 3484, and 7437 over-counting samples and 1951, 2073, 2425, and 466 under-counting samples for MODES, DOWA, thermal, and vibration sources, respectively. The total number of samples was 12078. In the following sections, we discuss the energy use, temperature effectiveness (for quality of comfort), and ventilation results obtained in our simulations.

5.1 Energy Use

In this section, we analyze the effect of each data scheme on energy consumption. The building's HVAC system comprises of a single duct terminal reheat composed of an Air Handler Unit (AHU) and Variable Air Volume (VAV) boxes. The AHU includes a fan, heating, and cooling coils that can change the air's temperature. The VAV boxes take this pre-conditioned air from the main duct and control the airflow provided to each zone. The power consumption sources include the supply fan, heating coils, and cooling coils. The HVAC control method is an EnergyPlus built-in rule-based control method based on occupied and unoccupied zone information. The heating and cooling setpoints are 21.1°C and 23.9°C in the working time (07:00 am - 06:00 pm) and 12.8°C and 40°C in the non-working time (6:00 pm - 07:00 am).

Figure 8 shows the monthly energy consumption in two scenarios for four studied data schemes (MODES, DOWA, Thermal, and Vibration). The EnergyPlus controller operates with occupancy information provided by each of the occupancy sensing schemes, with the over- and under-counting values to the ground truth for each scheme mentioned above. In the case of over-counting, we see in Figure 8-(a) that the energy use of the vibration-only scheme is very high since this scheme has the highest over-counting values. The HVAC system will tend to consume significantly more energy

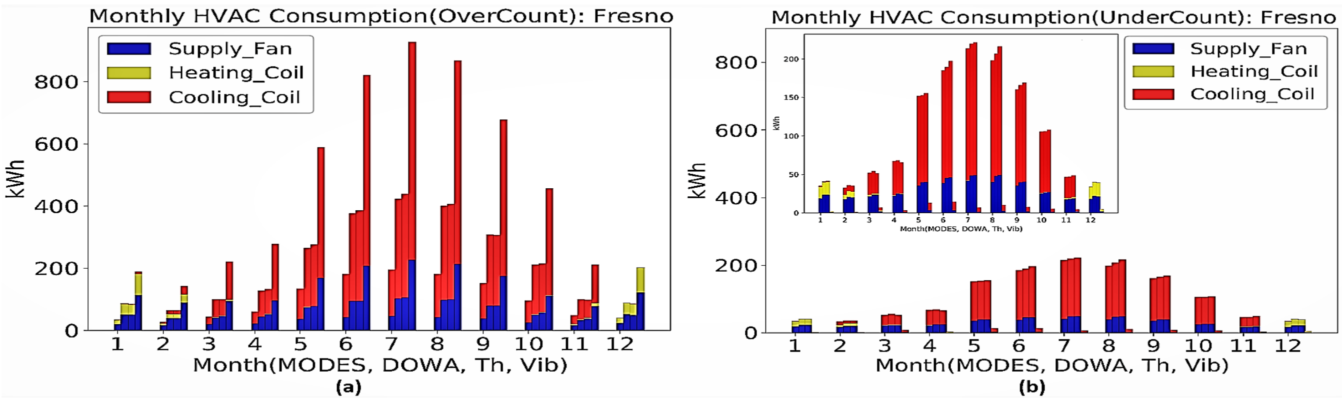

Figure 8: Monthly energy consumption (with zoom-in view) for (a) over-counting, and (b) under-counting scenarios.

by trying to condition zones that may be empty, but the vibration scheme informs that they are occupied. Of all the occupancy schemes used, we see that MODES provides the best energy results since it tends to over-count the least of all the schemes tested. In terms of under-counting, all the schemes have very similar energy use except for the vibration scheme. Note that by under-counting, the HVAC controller will tend to float the temperature in zones that it believes to be empty, even though they are occupied, so the quality of service will suffer as seen below.

5.2 Temperature Effectiveness

We studied the effects of our data schemes on the building temperature effectiveness, i.e. the ideal temperature that should be provided to the occupants for quality of comfort. To be ASHRAE [14] compliant, we must maintain the set-point temperatures to ensure that $-0.5 \leq PMV \leq 0.5$. PMV (Predictive Mean Vote) is calculated by Fanger's equation [25]. PMV predicts the mean thermal sensation vote on a standard scale for a large group of persons. The American Society of Heating Refrigerating and Air Conditioning Engineers (ASHRAE) developed the thermal comfort index by using coding -3 for cold, -2 for cool, -1 for slightly cool, 0 for natural, +1 for slightly warm, +2 for warm, and +3 for hot. PMV has been adopted by the ISO 7730 standard. The ISO recommends maintaining PMV at level 0 with a tolerance of 0.5 as the best thermal comfort. Fanger's PMV depends on temperature, humidity, air velocity, occupants' clothing, and activity. Based on the PMV equation, we get the best temperature when the PMV is 0. Then, we compare the ideal temperature with the temperature under different data schemes. For this analysis, we examine the root mean square error (RMSE) of the zone temperature difference per person between these two values.

Figure 9 shows the product of room temperature RMSE and the number of occupants for four data schemes under the over-counting and under-counting scenarios. In the over-counting scenario (a), the best quality of comfort (i.e. smallest RMSE deviation) is provided by the vibration scheme. However, this high quality of comfort comes with a high price tag, since this is achieved by over-conditioning the spaces and using a lot of energy as seen in the previous section. Of the remaining schemes, MODES produces the best quality of comfort (and the lowest energy use as seen before). Both DOWA and Thermal have very similar quality of comfort results. In the under-counting scenario, MODES produces the best quality of service of all the schemes (even with comparable energy use as seen before). We note that the vibration scheme produces the worst quality of comfort by a large margin. The energy savings presented in the previous section come at a significant cost in the quality of comfort, as the HVAC controller will save energy by not conditioning certain zones that are in reality occupied, trading off energy use for lower occupant comfort.

5.3 Ventilation

ASHRAE Standard 90.1 mandates the Demand-Controlled-Ventilation (DCV) system for densely occupied spaces since the 1999 version and also requires the DCV system to comply with ASHRAE Standard 62.1 (ASHRAE 2019). We use the formula in the standard to constantly calculate the minimum ventilation requirements based on the number of occupants in each zone, actuated by our HVAC controller. Figure 10 shows the simulation of annual CO_2 concentration (ppm) provided by each one of our data schemes in both over-counting and under-counting cases. In the over-counting scenario shown in Figure 10-(a), all schemes follow a similar pattern, with higher CO_2 concentrations in Summer and Winter, and lower in the Spring and Fall (shoulder seasons) when the controller takes advantage of the mild weather to use a lot of external air that is close to the temperature setpoints required. We notice that the vibration scheme delivers a slightly improved result compared to others since it is over-conditioning the space as shown above. In the under-counting scenario shown in Figure 10-(b), we see a similar case for all the schemes but the vibration scheme. This is because there is a higher chance of a zone not being ventilated properly due to incorrect (under-count) information provided by the vibration sensor. Also, please note that the MODES scheme results in a slightly lower concentration compared to the thermal and DOWA schemes.

6 DISCUSSION

The late-fusion occupancy estimation strategy provided by MODES could be used with other types of occupancy sensors. As it was shown in the paper, MODES should be able to address most of the deficiencies of its input modules, especially for the vibration sensor. However, the degraded performance of the vibration source in our work could cause significant accuracy degradation. One thing worth

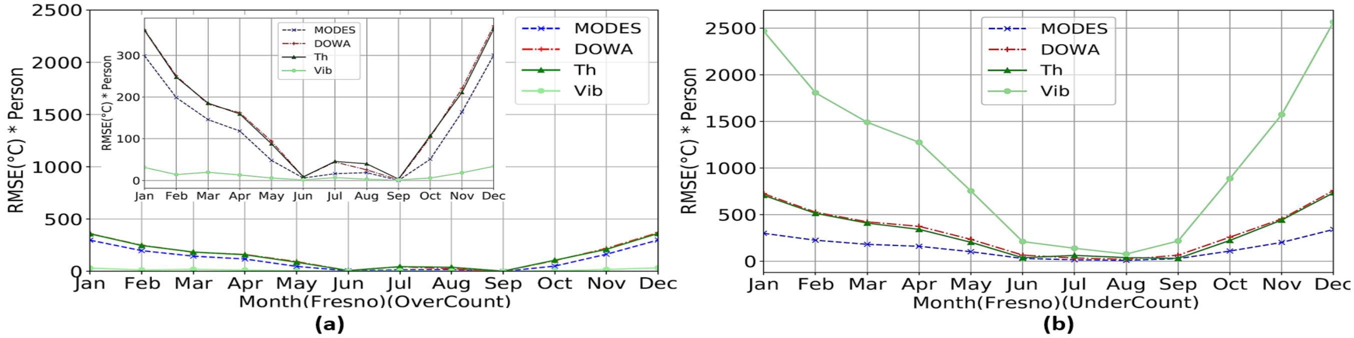

Figure 9: Monthly temperature RMSE for over-counted occupants for (a) over-counting (with its zoom-in view), and (b) under-counting scenarios.

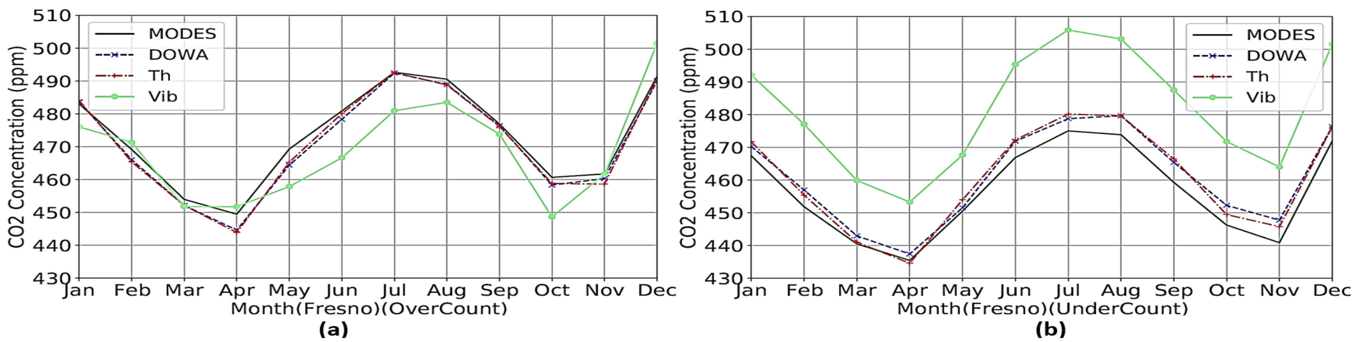

Figure 10: Monthly CO_2 concentration simulated by EnergyPlus for (a) over-counting, and (b) under-counting scenarios.

analyzing may be the level of sensor inaccuracies that MODES can correct in the limit. We leave this analysis for future work.

DOWA tries to optimally fuse the occupancy data between two input streams, constraining its output to select an optimal occupancy value inside the range of two inputs. However, there are limitations in further generalizing it to more than two sources. In this case, each source has to be assigned a quadratic weight function that can cause higher computational and training costs for MODES. The development and evaluation of MODES with more than two occupancy sources are left for future work.

Another point of concern is the amount of training data required. We have shown that for DOWA and MODES we need only less than a week of data to get a good estimate. However, the BMC occupancy model used in the particle filter requires a minimum of one week, and possibly more for more accurate results. In addition, occupancy patterns tend to change with the seasonality of the data, which means that the BMC model must be retrained as the occupancy patterns change over the years. We believe that one way to tackle this issue is to use the data provided by MODES to do the retraining. While the data is very accurate, it is not as accurate as the ground truth. However, it may be accurate enough to be able to capture the seasonality of the data as the occupancy patterns change. We leave a more detailed analysis of this issue for future work.

7 CONCLUSION

In this paper, we developed MODES; a Multi-sensor Occupancy Data-driven Estimation System for smart buildings. It applies a late (decision level) sensor fusion on two occupancy data streams collected and processed by thermal and vibration sensors, individually. The sensor fusion in MODES consists of a Data-driven Optimization based Weighted Average (DOWA) module that assigns more weights to the more reliable data stream in a non-linear way for different input values. To further improve the final occupancy estimate, we pass DOWA's output to a particle filter with a Blended Markov Chain (BMC) transition model. Our results show that MODES could achieve an accuracy of 0.73 and 0.84 in high and low occupancy scenarios. These outperform the state-of-the-art late fusion technique i.e. linear regression by 21% and 30%, the thermal input by 19% and 20%, and the vibration input by 40% and 33%, for high and low occupancy, respectively. Based on our analysis, about one week of training data would be sufficient to achieve optimal occupancy classification accuracy. Our EnergyPlus simulations show that MODES can save a significant amount of energy use both in Summer and Winter in a building.

8 ACKNOWLEDGEMENTS

We would like to thank anonymous reviewers for their constructive comments and helpful suggestions. This material is based upon work partially supported by a 2020 Seed Fund award from the Center of Information Technology in the Interest of Society (CITRIS) and the UC National Laboratory Fees Research Program of the University of California, grant# 69763.

REFERENCES

[1] [n.d.]. EnergyPlus. https://energyplus.net/.
[2] [n.d.]. How much energy is consumed in U.S. residential and commercial buildings? https://www.eia.gov/tools/faqs/faq.php?id=86&t=1.
[3] [n.d.]. ridge regression. https://doi.org/10.1093/oi/authority.20110803100420839
[4] Stefan Achleitner, Tao Liu, Ankur U. Kamthe, and Alberto E. Cerpa. 2013. Short Term Solar Energy Forecasting Using Wireless Sensor Networks. In *Proceedings of the Second UC Solar Research Symposium*.
[5] F. Adib, Z. Kabelac, and D. Katabi. 2015. Multi-person localization via RF body reflections. In *In Proceedings of the 12th USENIX Conference on Networked Systems Design and Implementation*. 279–292.
[6] Omid Ardakanian, Bhattacharya, and David Culler. 2016. Non-intrusive techniques for establishing occupancy related energy savings in commercial buildings. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*. ACM, 21–30.
[7] Alex Beltran, Varick L. Erickson, and Alberto E. Cerpa. 2013. ThermoSense: Occupancy Thermal Based Sensing for HVAC Control. In *Proceedings of the Fifth ACM Workshop on Embedded Sensing Systems for Energy-Efficient Buildings (BuildSys 2013)*. https://doi.org/10.1145/2528282.2528301, 1–8.
[8] N Cao, J Ting, S Sen, and A Raychowdhury. 2018. Smart sensing for HVAC Control: collaborative intelligence in optical and IR cameras. In *IEEE Trans. Ind. Electron.*, Vol. 12. IEEE, 443–455.
[9] Zhenghua Chen, Mustafa K.Masood, and Yeng Chai Soh. 2016. A fusion framework for occupancy estimation in office buildings based on environmental sensor data. *Energy and Buildings* 133 (2016), 790–798.
[10] Ken Christensen, Ryan Melfi, Bruce Nordman, Ben Rosenblum, and Raul Viera. 2013. Using existing network infrastructure to estimate building occupancy and control plugged-in devices in user workspaces. *International Journal of Communication Networks and Distributed Systems* 12, 1 (2013), 4–29.
[11] G. Conte and et a. 2014. a first approach using ibeacon for an energy efficient occupancy detection system. In *1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*. ACM, 11–19.
[12] A. Corna and et al. 2015. Occupancy detection via ibeacon on android devices for smart building management. In *2015 Design, Automation & Test in Europe Conference & Exhibition*. 629–632.
[13] Aveek K. Das and et al. 2017. Non-Intrusive Multi-Modal Estimation of Building Occupancy. In *SenSys '17: Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*. 399–402.
[14] Richard J. De Dear and Gail S. Braiger. 2002. Thermal comfort in naturally ventilated buildings: revisions to ASHRAE Standard 55. *Energy and Buildings* 34, 6 (2002), 549–561.
[15] Xianzhong Ding, Wan Du, and Alberto Cerpa. 2019. Octopus: Deep reinforcement learning for holistic smart building control. In *ACM BuildSys*.
[16] Xianzhong Ding, Wan Du, and Alberto E Cerpa. 2020. Mb2c: Model-based deep reinforcement learning for multi-zone building control. In *ACM BuildSys*.
[17] Robert H. Dodiera, Gregor P. Henzeb, Dale K. Tiller, and Xin Guo. 2006. Building occupancy detection through sensor belief networks. *Energy and Buildings* 38, 9 (2006), 1033–1043.
[18] Bing Don and et al. 2010. An information technology enabled sustainability test-bed (ITEST) for occupancy detection through an environmental sensing network. *Energy and Buildings* 42 (2010), 1038–1046.
[19] Tobore Ekwevugbe. 2013. *Advanced Occupancy Measurement Using Sensor Fusion.* Ph.D. Dissertation. Institute of Energy and Sustainable Development De Montfort University, Leicester, United Kingdom.
[20] Tobore Ekwevugbe and et al. 2017. Improved occupancy monitoring in non-domestic buildings. *Sustainable Cities and Society* 30 (2017), 97–107.
[21] Varick Erickson, Miguel A. Carreira-Perpinan, and Alberto E. Cerpa. 2011. OBSERVE: Occupancy-based system for efficient reduction of HVAC energy. In *Proceedings of the Tenth ACM/IEEE International Conference on Information Processing on Sensor Networks (IPSN 2011)*. ACM/IEEE, 258–269.
[22] Varick Erickson, Miguel Carreira-Perpiñán, and Alberto E Cerpa. 2011. OBSERVE: Occupancy-based system for efficient reduction of HVAC energy. In *Proceedings of the 10th ACM/IEEE international conference on information processing in sensor networks (IPSN)*. IEEE, 258–269.
[23] Varick Erickson and Alberto Cerpa. 2012. Thermovote: Participatory Sensing for Efficient Building HVAC Conditioning. In *Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficient Buildings (BuildSys 2012)*. ACM, 0–16.
[24] Varick Erickson and Alberto E Cerpa. 2010. Occupancy based demand response HVAC control strategy. In *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building (BuildSys 2010)*. ACM, 7–12.
[25] Povl O. Fanger. 1970. Thermal comfort. Analysis and applications in environmental engineering. *Copenhagen: Danish Technical Press.* (1970).
[26] G. Fierro and et al. 2012. Zone-level occupancy counting with existing infrastructure. In *Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*. ACM, 205–206.
[27] William J Fisk, David Faulkner, and Douglas P Sullivan. 2006. *Accuracy of CO_2 Sensors in Commercial Buildings: A Pilot Study.* Technical Report LBNL- 61862. Lawrence Berkeley National Lab.
[28] Christian Groß, Reuben Borrison, Johannes Schmitt, and Markus Aleksy. 2020. Towards an Occupancy Count Functionality for Smart Buildings - An Industrial Perspective. In *2020 2nd IEEE International Conference on Industrial Electronics for Sustainable Energy Systems (IESES)*, Vol. 1. 331–336. https://doi.org/10.1109/IESES45645.2020.9210641
[29] Brodie W. Hobsona and et al. 2019. Opportunistic occupancy-count estimation using sensor fusion: A case study. *Building and Environment* 159 (2019), 106154.
[30] Zhizhang Hu, Tong Yu, Yue Zhang, and Shijia Pan. 2020. Fine-grained activities recognition with coarse-grained labeled multi-modal data. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*. 644–649.
[31] Ltd. D&R International. [n.d.]. 2011 Building Energy Data Book. U.S. Dept. of Energy. https://openei.org/doe-opendata/dataset/buildings-energy-data-book.
[32] M. Jin, N. Bekiaris-Liberis, K. Weekly, NC.J. Spanos, and A.M. Bayen. 2002. Occupancy detection via environmental sensing. *IEEE Trans. Autom. Sci. Eng.* 34, 6 (2002), 549–561.
[33] Ankur Kamthe, Lun Jiang, Matthew Dudys, and Alberto Cerpa. 2009. SCOPES: Smart Cameras Object Position Estimation System. In *EWSN 2009: Wireless Sensor Networks*. European Conference on Wireless Sensor Networks, 279–295.
[34] Neil E Klepeis, William C Nelson, Wayne R Ott, John P Robinson, Andy M Tsang, Paul Switzer, Joseph V Behar, Stephen C Hern, and William H Engelmann. 2001. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *Journal of Exposure Science and Environmental Epidemiology* 11, 3 (2001), 231–252.
[35] Devanshu Kumar, Xianzhong Ding, Wan Du, and Alberto Cerpa. 2021. Building sensor fault detection and diagnostic system. In *ACM BuildSys*.
[36] Timilehin Labeodan, Wim Zeiler, Gert Boxem, and Yang Zhao. 2015. Occupancy measurement in commercial office buildings for demand-driven control applications—A survey and detection system evaluation. *Energy and Buildings* 93 (2015), 303–314.
[37] Ubaid Mehmood, I. Moser, Prem Prakash Jayaraman, and Abhik Banerjee. 2019. Occupancy Estimation using WiFi: A Case Study for Counting Passengers on Busses. In *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*.
[38] Ryan Melfi, Ben Rosenblum, Bruce Nordman, and Ken Christensen. 2011. Measuring building occupancy using existing network infrastructure. In *2011 International Green Computing Conference and Workshops*. DOI: 10.1109/IGCC.2011.6008560.
[39] Mostafa Mirshekari, Jonathon Fagert, Shijia Pan, Pei Zhang, and Hae Young Noh. 2021. Occupant Localization in Obstructive Indoor Environments Using Footstep-Induced Floor Vibrations. In *Dynamics of Civil Structures, Volume 2.* Springer, 121–124.
[40] Abdallah Naser, Ahmad Lotfi, and Junpei Zhong. 2021. Adaptive Thermal Sensor Array Placement for Human Segmentation and Occupancy Estimation. *Buildings and environments* 21, 2 (2021), 1993–2002.
[41] Abdallah Naser, Ahmad Lotfi, and Junpei Zhong. 2021. Occupancy determination based on time series of CO2 concentration, temperature and relative humidity. *Buildings and environments* 21, 2 (2021), 1993–2002.
[42] Shijia Pan, Amelie Bonde, Jie Jing, Lin Zhang, Pei Zhang, and Hae Young Noh. 2014. Boes: building occupancy estimation system using sparse ambient vibration monitoring. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2014*, Vol. 9061. International Society for Optics and Photonics, 90611O.
[43] Shijia Pan, Mostafa Mirshekari, Pei Zhang, and Hae Young Noh. 2016. Occupant traffic estimation through structural vibration sensing. In *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2016*, Vol. 9803. International Society for Optics and Photonics, 980306.
[44] Shijia Pan, Tong Yu, Mostafa Mirshekari, Jonathon Fagert, Amelie Bonde, Ole J Mengshoel, Hae Young Noh, and Pei Zhang. 2017. Footprintid: Indoor pedestrian identification through ambient structural vibration sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–31.
[45] Pratiksha Rashinkar and V. S. Krushnasamy. 2017. An overview of data fusion techniques. In *2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*. DOI: 10.1109/ICIMIA.2017.7975553.
[46] Cees G. M. Snoek and et al. 2019. Early versus late fusion in semantic video analysis. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*. 106154.
[47] Marina Sokolova, Nathalie Japkowicz, and Stan Szpakowicz. 2006. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In *AI 2006: AI 2006: Advances in Artificial Intelligence*, Vol. 4304. 1015–1021.
[48] Robert Tomastik, Yiqing Lin, and Andrzej Banaszuk. 2008. Video-based estimation of building occupancy during emergency egress. In *2008 American Control Conference*. IEEE.
[49] Ash Tyndall, Rachel Cardell-Oliver, and Adrian Keating. 2016. Occupancy Estimation Using a Low-Pixel Count Thermal Imager. *IEEE Sensors Journal* 16, 10

Table 3: Weight assignment to different features as presented in Groß et al. [28].

	Accuracy	Range	Coverage	Delay	Comp. Cost	Privacy	Ease of inst.	Cost
Accuracy	A	R	A	A	CC	P	A	A
Range		R	C	R	R	P	R	R
Coverage			C	D	CC	C	C	C
Delay				D	D	P	E	CO
Comp. Cost					CC	P	CC	CO
Privacy						P	P	P
Ease of inst.							E	CO
Cost								CO

(2016), 3784–3791.

[50] Daniel A Winkler, Ashish Yadav, Claudia Chitu, and Alberto E Cerpa. 2020. OFFICE: Optimization Framework For Improved Comfort & Efficiency. In *2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN).* ACM, 265–276.

[51] Beilei Xu, Orhan Bulan, Jayant Kumar, Safwan Wshah, Vladimir Kozitsky, and Peter Paul. 2015. Comparison of Early and Late Information Fusion for Multi-camera HOV Lane Enforcement. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems.* 913–918. https://doi.org/10.1109/ITSC.2015.153

[52] Junjing Yang, Mattheos Santamouris, and Lee Siew Eang. 2016. Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings. *Energy and Buildings* 121 (2016), 344–349.

[53] Hengyang Zhao, Qi Hua, Hai-Baq Chen, Yaoyao Ye, Hai Wang, Sheldon X.-D. Tan, and Esteban Tlelo-Cuautle. 2018. Thermal-Sensor-Based Occupancy Detection for Smart Buildings Using Machine-Learning Methods. *ACM Transactions on Design Automation of Electronic Systems* 22, 4 (2018), 1–21.

[54] Han Zoua, Yuxun Zhoua, Hao Jiangc, Szu-Cheng Chiend, Lihua Xie, and Costas J. Spanos. 2018. WinLight: A WiFi-based occupancy-driven lighting control system for smart building. *Energy and Buildings* 158 (2018), 924–938.

[55] M.S. Zuraimi, A. Pantazaras, K.A. Chaturvedi, J.J. Yang, K.W. Tham, and S.E. Lee. 2018. Predicting occupancy counts using physical and statistical Co2-based modeling methodologies. *Buildings and environments* 123, 2017 (2018), 517–528.

A SENSOR SELECTION

We have included in this Appendix the main features used to select the sensor technologies used in our work. Groß et al. in [28] presented a Pugh Matrix providing a visual comparison between individual state-of-the-art occupancy sensing techniques based on a set of features and requirements. Each feature in the matrix is assigned to pick a rating of -1, 0, or 1, in the order of the worst-to-best state. Those rating values were then summed up to deliver an evaluation metric for each occupancy estimation technique. The features evaluated were as follows:

- **Accuracy:** The ratio of true positives/negatives in an occupancy estimation process.
- **Detection Range:** The total range of occupancy counting by a particular technology.
- **Sensor Coverage Area:** Coverage is generally referred as the maximum possible distance in which a sensor could detect occupants. We considered the blockage and NLOS issues to be included in this requirement, as well.
- **Delay:** It refers to the sensor's delay in occupancy detection.
- **Computational costs:** It refers to the amount of computational complexity required by the method to capture, process, and/or transmit the occupancy data.
- **User's Privacy concerns:** As to what level of privacy information a user is willing to share.
- **Ease of installation:** The amount of manual effort/time and knowledge needed to deploy and set up an occupancy estimation system.
- **Cost:** It refers to the hardware and installation costs of the occupancy estimation system.

The authors in [28] assigned a weight to each feature, based on the importance of each one. Table 3 shows a confusion matrix through which our features are competing against each other. For example, in case the Accuracy (A) is considered to be more important than the Coverage (C), the associated entry in the matrix is filled by an A. The weight for a specific feature is then calculated by counting the total number of entries containing that feature. Finally, the weighted sum is calculated and entered in Table 4 for each occupancy estimation technology.

According to the comparison in Table 4, the most two promising techniques are the Heat-map (i.e., thermal-based) and Vibration sensing systems. Therefore, we selected these two state-of-the-art technologies for our work. We picked them since they have complementary strengths. More details can be found in [28].

Table 4: Pugh Matrix to compare different individual occupancy estimation technologies as presented in Groß et al. [28].

	weight	CO_2	Video-based	WiFi	Heat map	Vibration	PIR	Ultrasonic	LIDAR	Humidity	RF signal	Power meter
Accuracy	5	-1	1	0	0	0	-1	0	1	-1	-1	-1
Range	6	0	1	0	0	0	-1	-1	1	0	0	0
Coverage	5	-1	0	0	0	1	0	0	0	-1	1	-1
Delay	3	-1	1	0	1	1	1	1	1	-1	1	-1
Computational Cost	4	1	-1	-1	1	0	1	1	0	1	0	1
Privacy	7	1	-1	0	1	1	1	1	0	1	0	1
Ease of installation	2	0	0	-1	0	-1	1	1	0	0	-1	1
Cost	4	1	0	1	1	0	1	0	-1	1	0	1
Weighted Sum		2	3	-2	18	13	9	10	10	2	1	4