Learning occupants’ workplace interactions from wearable and stationary ambient sensing systems

Ali Ghahramani\textsuperscript{a,b}, Jovan Pantelic\textsuperscript{b}, Casey Lindberg\textsuperscript{b}, Matthias Mehl\textsuperscript{c}, Karthik Srinivasan\textsuperscript{d}, Brian Gilligan\textsuperscript{e}, Edward Arens\textsuperscript{a}, for the Wellbuilt for Wellbeing Project Team\textsuperscript{1}

\textsuperscript{a} Center for the Built Environment, University of California, Berkeley, 390 Wurster Hall, Berkeley, CA 94720, USA
\textsuperscript{b} Institute on Place and Wellbeing, University of Arizona College of Medicine, P.O. Box 245153, Tucson, AZ 85724, USA
\textsuperscript{c} Department of Psychology, University of Arizona, 1503 E University Blvd., Tucson, AZ 85721, USA
\textsuperscript{d} INSITE Center for Business Intelligence and Analytics, Department of MIS, Eller College of Management, University of Arizona, 430 McClelland Hall, Tucson, AZ 85721, USA
\textsuperscript{e} General Services Administration, Washington DC, USA

HIGHLIGHTS

- A framework to learn occupant interactions from ambient sensing technologies is proposed.
- Several machine learning algorithms are tested and the one that outperforms others is selected.
- The framework selects proper the technologies and averaging windows.
- 221 employees of federal agencies participated in this study.
- Using random forest algorithm an accuracy of 86.72% to predict interactions was obtained.

ABSTRACT

Having access to real-time information on building occupants’ state of interactions enables optimization of building systems for improved energy efficiency, well-being and productivity of the occupants. In this paper, we propose a framework to learn occupant interactions from ambient sensing technologies (e.g., sensing of variables such as sound (dB), CO\textsubscript{2} (ppm), light intensity (lux), dry-bulb temperature (°C), relative humidity (RH%), pressure (mbar)) from both stationary and wearable devices and select the technologies and averaging windows which contain the required information for learning. In this framework, several supervised machine learning algorithms are tested on the labeled datasets and the algorithm which outperforms others is selected. Two types of sensing devices were utilized for analyses: wearable devices worn around the neck by the test subjects, and a network of stationary devices located in the test subjects' working indoor spaces. 221 employees of federal agencies housed in facilities managed by the US. General Services Administration in the mid-Atlantic and Southern states participated in this study, answering questions about their current task every hour. Overall accuracies were observed of 86.72% for wearable and stationary devices, 81.25% for only wearable-only, and 85.16% for stationary-only for prediction of the mixed multi-label classification via Random Forests algorithm. The high prediction allows for identifying subjects’ tasks when training labels are not available. Predicting occupants’ interactions as a main indicator of occupants’ behavior have significant implications for the energy efficiency of building systems (up to 20% savings).

1. Introduction

In the built environment, an occupant interacts physically with building elements and socially with other occupants. Both of these interactions have implications for different types of efficiencies (e.g., building energy use and organization workflow). Occupants interact...
physically with the building in adjusting environmental conditions (e.g., lighting level, temperature) according to their preference, and in activating appliances (e.g., plug loads) necessary for a service [1]. Social interactions among occupants also impact building systems performance through variations in space utilization and the energy necessary to keep the indoor environment conditioned. Buildings account for approximately 40% of the total energy consumed in the United States [2]. Previous research has shown that optimizing building systems performance based on occupant behavior and interactions can reduce overall building energy consumption up to 20% [3,4]. In addition, energy required to service unoccupied spaces may account for almost 50% energy used in buildings [5], which could be reduced if occupancy were monitored through real-time occupant workplace interactions. The International Energy Agency has also emphasized the monitoring of the occupants’ behavior and specifically interactions as a tool for more efficient building systems [6]. Beyond energy issues, interactions between people in work spaces have been shown to affect productivity within organizations [7]. Variations in patterns of communication may account for up to 40% variation in productivity. Occupants’ behavior is shaped by social circuits (formed by the back-and-forth pattern of signaling between people, unconscious reflexes evolved for social coordination) that work to fuse us together into a coordinated whole [8]. Considering the importance of topics like building energy consumption, building space utilization, and productivity research has begun to focus on learning the state of interactions among occupants in buildings. Given a real-time access to the interactions state among occupants, building energy and workplace efficiency can be improved by better allocating resources and optimizing systems performance. Information on interactions among occupants can determine personal spaces which are unoccupied and require less services. A real-time feed of the required services to the building management systems can for example issue on/off system commands to appliances or lighting systems or change temperature setpoints in the HVAC system.

However, most information about interactions among building occupants is currently obtained from occupants’ direct feedback (i.e., self-reporting), which is error-prone and may not provide information in a timely manner. In addition, sustaining such participatory involvement for long periods of time is challenging due to survey fatigue. To address these challenges, environmental sensing has been used to indirectly capture human interactions by tracking human social signatures in the indoor environments. Due to advancements in sensing technologies and reductions in sensor costs, indoor and personal sensing arrays have become commercially feasible. Collected environmental signatures can also be used for further understanding the environmental quality and human interactions that occur within the built environment. However, research in the literature thus far has been limited to small case-studies with few subjects and sensing technologies. No large-scale and long-duration studies have been performed monitoring occupants’ workplace interactions. In addition, there is a gap in the literature about methods for selecting machine learning techniques and sensing technologies among a group of potential ones in a hierarchical order for predicting occupants’ workplace interactions.

In this study, we have proposed and validated a framework for predicting occupants’ workplace interactions as defined by an organization (e.g., U.S. General Services Administration (GSA)) based on stationary and wearable sensing arrays. The framework also uses a recursive feature elimination method to select sensing technologies that provide the optimal information gains and prevent overfitting. Major performance measures for the interaction modeling are introduced and the advantages of each one is discussed. Due to the fact that short term measurements of sensors might not correctly capture the dynamics of ambient conditions as they relate to the interactions, we included a component in the framework to generate different sliding window lengths and select the optimal values. In cases where labels are mutually exclusive and could be coupled for a more in depth understanding of interactions in an environment, the framework utilizes a multi-label classification reformulation. The data collection was conducted in facilities managed by GSA in the mid-Atlantic and Southern states. 221 employees of federal agencies participated in this study and were asked about their current task every hour. Ambient sensing technologies (e.g., sensing of variables such as sound (dB), CO2 (ppm), light intensity (lux), dry-bulb temperature (°C), relative humidity (RH %), pressure (mbar)) were collected from wearable devices worn around the neck by the test subjects, and from a network of stationary devices located in the test subjects’ working spaces. After cleaning the dataset, the framework was validated based on the described performance measures.

The structure of the paper is as follows. Section 2 provides a review of recent studies of the occupants’ interaction learning and modeling, and the gap that this study addresses. In Section 3, we describe our proposed methodology for modeling the interactions, selecting sensing technologies, and optimizing the performance by using moving averages and labels refinement. In Section 4, the data collection set up and procedures done by GSA which paved the path for our explorations are discussed. We present the results of our methodology in Section 5. In Section 6, we provide the energy savings opportunities from control schedule modifications based on occupants’ workplace interactions real-time feed. Section 7 provides a discussion on the generalization of the results and Section 8 describes the limitations and future steps of the study. Finally, Section 9 summarizes the results and concludes the paper.

2. Literature review

Direct occupants’ feedback via an online or offline survey or questionnaire is the most common technique for assessing occupants’ interactions in indoor environments. This technique is subject to reporting errors by occupants. In addition, lack of such information from occupants ultimately results in inefficiencies in building systems and excessive operation costs due to conservative operation choices made when real-time feedback is not available. Research has shown that filling out surveys over extended periods of time experiences reduction in participation by the occupants. Previous research has shown that temporal patterns in multiple streams of sensor data could help automatic analysis of human behaviors and habits in the ambient environment [9]. Specifically, for monitoring occupants interactions, Goughbury [10] used a wearable device having infrared transceiver, a microphone, and two accelerometers and applied dynamic Bayesian Networks to model interactions for a group of eight researchers for a few days. Various techniques for interpreting resident behavior patterns and determining when multiple residents are interacting based on sensors data such as power consumption meters and contact sensors were used. The effectiveness of their techniques was evaluated using two physical smart environment test beds. The performance accuracy in detecting conversations was 63.5% overall and 87.5% for conversations greater or equal to one minute. To account for the small sensor variability and analysis methods, Pentland [11] describes analyzing wearable sensor data with several statistical learning methods such as principal components analysis and clustering methods in order to make reliable estimates of users’ interactions. He presents a detailed description of eigenbehavior modeling for learning and classifying user behavior from proximity and location data, and influence modeling for predicting the behavior of a subject from another subject’s data. A binary decision boundary at 0.45 produces an equal-error accuracy of 87% of prediction. Later, Paradiso et al. [12] demonstrated the use of electronic badges as a tool that aids social interactions in the large conference events using large LED display, wireless infrared and radio frequency networking, and a host of sensors to collect data. Without taking into account personal characteristics, history, or other prior knowledge over 80% accuracy of personal interactions was realized. However, all these methods focus on a small group of test subjects and sensors are often preselected for learning based on the expert...
knowledge. In a more detailed sensing technique, Olguín [13] used Bluetooth and IR sensors to detect interactions such as proximity and face-to-face time and nonlinear regression had a correlation coefficient of $r = 0.62$ with $p = 0.01$ (explaining about 30% of the variance in group interaction satisfaction). Similarly, Cattuto et al. [14] used Active Radio Frequency Identification (RFID) devices that assess mutual proximity in a distributed fashion by exchanging low-power radio packets. They analyzed the dynamics of person-to-person interaction (defined as face to face conversations) networks obtained in three high-resolution experiments carried out at different orders of magnitude in community size. The data sets exhibit common statistical properties and lack a characteristic time scale from 20 s to several hours. Using a slide window of size 20 s, their experimental framework allows the prediction of face-to-face interactions under certain pre-defined circumstances with accuracy of over 99%. However, such techniques have a very limited application for the built environment as it only covers face to face and close-range communications. Specifically, for elderly applications, an ear worn sensor (e-AR sensor) consisting of a 3-axis accelerometer, as well as a Pulse Oximeter for the measurement of oxygen saturation levels (SpO2) and heart rate was used by Atallah et al. [15] to model elderly people activities. Even though the results were promising, no accuracy result was reported and the costs of the sensing technologies and the small group of the test subjects limit the broader reliability analysis of the technique.

In summary, research efforts in the literature have often focused on a small number of sensing nodes, short durations of data collections, low number of participants, and very limited type of interactions. Short duration experimental procedures and small numbers of test subjects result in learning algorithms that overfit to specific environmental characteristics and thus would have problems for generalizability. Having access to both sensing measurements very close to the occupants and relatively far would also shed on how occupants’ activities can result in sensor measurement reading variations. In this study, we explored the applicability of simultaneously using rich wearable and stationary sensing devices on a large-scale data acquisition (221 office workers) to improve the accuracy of interactions prediction. Findings from such study would shed light on potential generalizability of patterns in interactions among building occupants on a large scale which would then help development of permanent sensing modules for creating a real-time feed to the building controls for improved efficiency and service optimization.

3. Methodology

In this paper, we propose a framework to learn interactions from stationary and wearable ambient sensing technologies (e.g., measuring sound, CO2, light intensity). The learning module would then predict state of interactions when labels (e.g., occupants feedback) are not provided, enabling a real-time feed for building controls. The schematic diagram of the framework is demonstrated in Fig. 1.

The first step in this framework is to set up sensing devices with potentially useful sensors for capturing interaction signatures in the ambient environment. Interactions among office workers might impact ambient conditions in the close proximity and also the background conditions. To measure the variations in ambient conditions, we selected six types of sensing technologies: (1) sound, (2) CO2, (3) light intensity, (4) temperature, (5) relative humidity, and (6) pressure. We then set up sensing nodes (including all mentioned sensing technologies) both in a small sensor box mounted around subject’s neck (to capture close proximity ambient conditions) and in a stationary box located in the subjects working spaces (to capture background ambient conditions). Another component in the proposed framework is the personal voting (personal, in-the-moment self-reporting of one’s activity) which happens randomly over the course of time [16]. The random sampling helps labeling the sensor measurements. In practice, the collected subjective report would be matched with the timewise closest measurement point from all sensors. To select the supervised learning method for feature selection algorithm, we compare various supervised machine learning algorithms (e.g., logistic regression, decision tree, K-nearest neighbor, Support Vector Machines (linear basis function), and Random Forests [17]). To implement the comparison, we first randomly separate rows of the data set into training (80%) and testing (20%) sets. The percentages are arbitrary and can be selected based on the data availability. For the training process, we use 10-fold cross-validation to tune hyperparameters of each model. In other words, each training set is divided into 10 parts and each time (10 times total), and the learning model is tuned based on the 9 parts and was validated based on the other 1 part. When training process is completed, we use the trained models to predict the state of interactions in the testing data set. The algorithm that outperforms other algorithms is then selected for the feature selection step. The labeled data set would then be fed into the feature selection component which selects features (i.e., sensing technologies) which provide the highest acceptable accuracy with smallest loss of information. The main reasons for implementing the feature selection on the sensing technologies are as follows: (1) simplification of models and reducing capital costs of sensor deployment, (2) shorter algorithm training computation expense, (3) avoiding the curse of dimensionality, and (4) reducing overfitting by eliminating unrelated features. There are various methods to implement feature selection. These approaches are often categorized into 3 groups: (1) filter method, (2) wrapper method, (3) embedded method. We opt to choose wrapper method because even though it is more computationally expensive than embedded method, it considers all permutations of the features to be selected in an unsupervised learning algorithm and unlike filter method it considers the inter-relationships between the features. In wrapper strategy (Fig. 2), we use a recursive feature elimination method which uses resampling to reduce bias.

In this method, we first divide the data set into training and test sets (to factor in the variability caused by feature selection when calculating performance) via a 10-fold cross validation approach and in each iteration fit the selected supervised learning algorithm to training set and use the generated model to calculate the importance of each feature. Let $S$ be the ordered sequence of the features numbers in terms of their importance ($S_1 > S_2 \ldots$). At each iteration of feature selection, the $S_i$ top ranked predictors are selected, the model is refit and performance is assessed. The value of $S_i$ with the best performance is determined and the top $S_i$ predictors are used to fit the final model [18]. We finally compare the results from all feature domain vs only wrapper selected features and evaluated the performance of the proposed framework.

3.1. Performance measures of training and final evaluations of the model

In order to tune hyperparameters of each classification learning model (e.g., weights of features in logit function of logistic regression algorithm), a performance measure to evaluate the goodness of fit of the models is required. The common practice is often to use the accuracy measure $\frac{TP + TN}{TP + TN + FP + FN}$ which gives equal values to all correctly and incorrectly predicted positives and negatives. However, in cases where the data set is slightly skewed towards positives or negatives (the ratio of positives/negatives is relatively large), other metrics for evaluation that reduce the bias towards a certain label need to be implemented. In our case, we use Kappa performance measure to reduce the bias of the learning towards higher frequency of a label. Cohen’s kappa coefficient ($\kappa$) (i.e., kappa measure) is a measure of inter-rater agreement for classification problems with reduction in elemental bias. $\kappa = \frac{p_o - p_e}{1 - p_e}$ where $p_o$ is the accuracy $\frac{TP + TN}{TP + TN + FP + FN}$ and $p_e$ is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of observer randomly getting each label. $p_e = \frac{1}{N} \sum_k n_k/n_k$, where for each label $k$, $N$ is the number of data rows, and $n_k$ is the number of times supervised learning algorithm predicts...
the label as $i$.

Due to complexities associated with interpreting the absolute values of Kappa and to better assess the performance measure for evaluation of each model for comparison used in this study were accuracy \( \frac{TP + TN}{TP + TN + FP + FN} \), recall \( \frac{TP}{TP + FN} \), and precision \( \frac{TP}{TP + FP} \). Accuracy demonstrates the overall degree of closeness of predictions with the actual labels. However, in the context of predicting interactions, we do not value both types of errors (FP and FN) similarly. FPs (false positive prediction of a state of interaction which is not true) is an error which we intend to minimize, while having some FNs (false negative prediction of a state of interaction is not true) is not a big issue. In other words, there is a trade-off between precision and recall. In algorithm tuning and selection in the context of interactions, maximizing precision while having recall in an acceptable range is the goal. Consequently, we demonstrate tuned-algorithm performance based on the Kappa measure, and compare it based on accuracy, recall, and precision. Based on the comparison, we can then choose which environmental sensing technologies to select and also the prediction model enables estimation of interactions state when the actual state via subjective feedback is not provided.

### 3.2. Moving average for improved accuracy

During the initial exploratory analysis of sensor measurements, we realized there are many cases where instantaneous sensor readings (specifically for the wearable device) might not correctly represent the ambient conditions since occupant activity can chaotically impact measurements (e.g., occupant exhaling on the wearable sensing device). To address this challenge, we also calculate a moving average of sensor measurements with different window lengths and implement the first stage selected supervised learning algorithm on the data to find the best window length. The window length depending the frequency of sensor measurements can hold considerably different values. If the measurement frequency is 1 Hz, the window length can reach up to 500 points which corresponds to 500 different training and testing cases and can become a very computationally expensive task. In order to address challenge, different search and metaheuristic algorithms such as genetic algorithms or simulated annealing can be used. In our case, since the measurements were already a moving average of 5 min, we only considered 3 more windows (i.e., 10, 15, and 20-min spans).

---

**Fig. 1.** Proposed framework for capturing interactions based on wearable and stationary sensing devices.

**Fig. 2.** Recursive feature elimination method to select features.

```plaintext
for Each Iteration in Cross-Validation do
    Partition data into training and test via cross-validation
    Train the supervised learning model based on the whole feature vector
    Predict labels for the test set
    Calculate variable importance or rankings
    for Each subset size $S_i$ $i=1 \ldots S$ do
        Keep the most important features
        Train the model on the training set using $S_i$ features
        Predict labels for the test set
    end
end
Calculate the performance measure over the $S_i$ using test set
Determine the number of features
Estimate the final list of features to keep in final model
Fit the final based on the optimal $S_i$ using the complete data set
```
4. Data collection

4.1. Participants and setting

Part of the US General Services Administration’s environmental sensing project, Wellbuilt for Wellbeing, office workers in a variety of roles were recruited in four Federal buildings for an observational study. Employees in areas of each building that housed organizations with approval from management were able to volunteer. A total of 248 expressed interest in participating in the study. Due to scheduling problems, sickness, and exclusionary criteria, 17 office workers did not participate, resulting in a total enrollment of 231 participants. They were enrolled serially from May 2015 to August 2016. Due to un-expected changes in work schedules, 8 of the 231 participants were only observed for two work days, rather than the full three. A total of 221 participants had valid work-time data collected. Participants were observed for three consecutive working days at the office, however 8 participants were only able to be observed for two days.

4.2. Environmental sensing

Environmental sensing devices (Aclima Inc., San Francisco, United States), were deployed for data collection. In areas of each building where participants’ workstations were located, networks of mounted, stationary sensors were distributed. These stationary sensors were equipped with environmental sensor modules to collect sound (dB), CO2 (ppm), light intensity (lux), dry-bulb temperature (°C), relative humidity (RH%), pressure (mbar). Individual stationary devices were located to sample average conditions within conditioned spaces and to avoid externalities that may affect the measurements. The devices were mounted approximately 1.5 m off of the ground, located at least 0.5 m from corners, windows, and doors, not directly in front of fans or heaters, not in direct sunlight, and located at least 1 m from equipment such as photocopiers and printers. Sensors were calibrated in ambient air at Aclima, Inc. prior to deployment at the study sites. These devices were populated with sensors and collocated with EPA-approved or similar quality reference equipment.

In addition to the stationary devices, participants wore a personal wearable environmental device while at the office. This device was worn on the chest suspended on a lanyard around the neck and contained the same set of sensor modules listed above for the stationary device.

Wearable devices were calibrated alongside stationary nodes, using the same reference equipment and quantitative analyses. Fig. 3 shows the stationary and wearable sensing devices.

For all sensors, raw sensor output was converted to corrected environmental measures via Aclima-specific sensor models. Each sensor was assigned unique model parameters during calibration. Data from the wearable devices were downloaded after each use and post-processed. Temperature, relative humidity, pressure, sound, and light data were collected at 1 s intervals and aggregated to 5-min intervals. CO2 data were collected at 17 s intervals and aggregated to 5-min intervals.

4.3. Ecological momentary assessment and location information

Throughout each work day of observation, as participants were performing their typical work duties, ecological momentary assessments (EMA) of current work situations were taken. Once an hour, at random times and no less than 30 min apart, participants were prompted by their smartphone to provide their current location within their building, and to select what they were currently doing from a checklist of possible activities. The movisensXS application (movisens GmbH, Karlsruhe, Germany) was used for this EMA data collection.

Current location data were used to tie environmental data from the nearest stationary device to each participant, and thus each personal, wearable device. Stationary devices that were within approximately 10 m and in the same conditioned space as the listed location were manually coded on the floor plan. Participants were assumed to be in the same location for the previous in average 2.5 min and the subsequent 2.5 min to their actual EMA prompt to match with sensor measurements (sensor measurements happened every 5 min). In order to increase the amount of known locations and thus the environmental data from the stationary devices, location information for a subset of participants was augmented by collecting calendar information and written logs.

In order to understand which activities workers were engaged in at the moment of their EMA prompt, participants selected what they doing at the moment of the prompt from a checklist. Each participant saw one of two possible versions of the prompt, shown in Fig. 4, but all participants had the ability to select ‘Working alone,’ ‘Working with other people,’ and ‘Talking informally with a colleague.’ Item Selection was not mutually exclusive, that is, participants could select multiple items from the list.
We temporally matched environmental sensing measurements with EMA responses and cleaned the data by removing rows with missing values resulting in 1390 complete rows of 5-min data across all participants. Out of the 1390 rows, there were 693 positive cases for working alone, 279 positive cases of working with other people, 90 positive cases of talking informally, and 12 cases of being in a conference call, 16 positive cases on being on a phone call, 8 positive cases of lunch, and 89 positive cases of being on a telephone call (or conference call). Due to the small ratio of positives over total responses for all interactions (except for the working alone and working with other people activities), we opted to focus on modeling and learning the two activities (e.g., working alone and working with other people) as we have an unbiased dataset for them (with no differences between the two user interfaces regarding these activities). We leave the analysis of the rest of the interactions to a future study.

4.4. Multi-label classification reformulation

In many cases (like our data collection procedure), the subjects’ responses to multiple choice questions or a group of binary questions can be reshaped into a new mutually exclusive single dimension to enhance learning from the data. For example, if asked “what is your current task” and there are 4 choices of response, one can treat the results of each choice as a binary variable and predict the performance of the prediction algorithm accordingly. It is logically and mathematically sound, but in cases where these choices are mutually inclusive, only one positive state could be accepted. For avoiding a secondary classifier logic in such cases, a linear transformation of multiple vectors of responses provides in a single variable with multi classes which are mutually exclusive can help simultaneous prediction of the labels. In our case, the new vector will have three distinct labels: (1) working alone, (2) working with other people, and (3) others. We also removed the “others” label category to help improving the accuracy of the prediction due to mutual inclusivity of the options in the category. In the Results Section, we provide the performance measures for the prediction of newly reformed dimension.

5. Results

As described in the proposed framework, after matching the sensor measurements with the subjective votes and cleaning the data, we trained different supervised learning methods based on a randomly selected training set. We used Kappa performance to reduce the bias towards the abundance of a certain label, but since Kappa’s absolute values are difficult to understand, we demonstrate accuracy, recall, and precision metrics. Table 1 and 2 summarizes the results for “working alone” and “working with other people” binary classification cases. To get a better understanding of how a wearable device and a stationary device would solely perform, we also trained the model for one each time. Therefore, in cases where the measurements of a wearable or stationary device is not available, the trained model can be used for prediction.

As it can be seen in Table 1 and 2, the superior performance of a fairly complex Random Forests algorithm demonstrates a relatively high non-linearity in the model and can be traced back to the nature of the activities. In other cases, relatively good performance of the algorithms demonstrates that the proposed sensing devices contained the required features for classification. Random Forests algorithm realized overall accuracies of 80.13% and 86.75% for predicting the two types of interactions.

Recall (the rate of labeling a prediction as positive) was 91.30% for working alone, but only 47.22% for working with other people. However, precision (accurate prediction of positives) was more balanced as it was 79.25% for working alone and 94.44% for working with other people. Note the fact that the accuracies are lower in cases that one device is being used, the decision to use the prediction results require stakeholders to define an acceptability threshold. In addition, it is interesting to note that a device that is being carry around provides lower accuracies for classifying interactions than a device that is stationary and located further away from you. It is likely due to the reason that the information from the wearable device carries information that can easily lead to confusion (e.g., CO2; sound) with other personal behaviors.

The next step in our framework is the feature selection component. Figs. 5 and 6 demonstrate the accuracy of each number of the features (i.e., number of ambient sensing elements) for the selected supervised learning algorithm (i.e., random forests). Table 3 demonstrates the ranking of the features in terms how much information gain (importance) we had for each task prediction.

As it can be seen in Figs. 5 and 6, the number of features which provides the maximum accuracy (filled blue square as the highest point on the accuracy line) for detecting both working alone and working with other people is 8. The rest of the features (starred in Table 3: Wear_relative_humidity, Wear_light, Wear_temperature, Wear_CO2) results in overfitting in the training set and are actually not useful at all to keep in the feature vector to train supervised learning models. Similarly, in both Figs. 5 and 6, the slope of the curve significantly drops after the fourth feature and the accuracy improvements are minor. The top four ambient sensing technologies can then be generalized as (Wall_sound, Wear_sound, Wall_CO2, Wall_pressure). Based on the recursive feature selection algorithm results, the decision makers at this point can decide which sensing technologies they would like to remove the sensing devices, and if they accuracies are not acceptable, further discussions on potential helpful sensing technologies can be made.

In order to further enhance the learning performance of the proposed learning methods, as explained in the methodology section, we apply moving average based on different sliding window lengths. We tested the instantaneous values (window length of 0 corresponding to
the 5 min average) and window length of 1 (10 min average), 2 (15 min average), and 3 (20 min average). The optimal values are shown in bold fonts.

Based on Table 4, the optimal window length in terms of the accuracy for both wearable and stationary sensing devices is 0 and it happens for two types of target activities. In other words, the instantaneous measurements of sensors provide the highest accuracy of prediction. The same holds true for “working with other people” and “working alone” based on the stationary device. In case of learning based on wearable device only, for both types of activities, the optimal accuracy happened at window length of 1 (i.e., 10 min average).

The final step is multi-label classification reformulation. Leaving out any other choice, and choosing among only “working alone” and “working with other people”.

As it can be seen in Fig. 7, we observe an overall accuracy of above 80% with only 4 features and the addition of features do not help accuracy improvement considerably. Ranking of features was: Wall_sound, Wear_sound, Wall_CO2, Wall_pressure, Wall_relative_humidity, Wall_light, Wear_relative_humidity, Wear_light, Wall_temperature, Wear_temperature, Wear_CO2 and they are very similar to Table 3 findings.

In Table 5, we have summarized the performance measures of the selected classification algorithm (i.e., Random Forests) for modeling the multi-label learning problem. The improvements in performance measures compared to Table 2 comes from the fact these labels are now mutually exclusive. In others, a person can either be working alone or working with other people and nothing else. This way the negative cases of both labels (i.e., not working alone and not working with other people) is removed and thus the performance measures are improved.

In summary, the very high prediction accuracy (above 80%) for all different types of sensing devices suggest that the selected ambient sensing technologies with the sample rates are reasonably capturing the state of the office worker’s interactions and thus can be used for the prediction of the cases where the labels (i.e., actual votes for the

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Wearable and Stationary Devices</th>
<th>Wearable Device</th>
<th>Stationary Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic regression</td>
<td>0.702 0.7935 0.7374</td>
<td>0.7152</td>
<td>0.8587 0.7248</td>
</tr>
<tr>
<td>K-nearest neighbors</td>
<td>0.6954 0.8587 0.7054</td>
<td>0.6887</td>
<td>0.8043 0.7184</td>
</tr>
<tr>
<td>Decision tree</td>
<td>0.755 0.9239 0.7391</td>
<td>0.7152</td>
<td>0.7935 0.7526</td>
</tr>
<tr>
<td>Support Vector Machines</td>
<td>0.7152 0.8587 0.7248</td>
<td>0.7351</td>
<td>0.837 0.7549</td>
</tr>
<tr>
<td>Random Forests</td>
<td>0.8013 0.913 0.7925</td>
<td>0.7483</td>
<td>0.8043 0.7872</td>
</tr>
</tbody>
</table>

**Table 1**

“Working alone” prediction performance for different algorithms.

**Table 2**

“Working with other people” prediction performance for different algorithms.
workers) are not available and only sensors measurements are available. Other performance measures such as recall and precision are also all relatively high and suggest the performance of the proposed sensing and learning method is accurate at calling out interactions and has a high accuracy of correct calls.

6. Energy savings from occupants’ workplace interactions feed

As explained earlier, previous research has shown that reducing energy wasted due to conservative operations ("always-on" mode during occupied hours) result in significant energy savings (up to 50%) [3–5]. In Fig. 8, we demonstrated the block diagram of information flow for optimizing the building systems based on the occupants’ interactions. Groups of appliances (i.e., HVAC systems, lighting systems, personal comfort devices, and appliances) can each be optimized by adjusting their control schedules and consequences according to occupants’ behavior and state of interaction in an office workplace. For lighting and the majority of office appliances, an on/off control logic for un-used systems would suffice to reduce energy consumption. It is slightly more difficult to reduce energy consumption for central HVAC systems since the systems have long response delays, but a control strategy that takes into account occupants’ thermal comfort profiles coupled with the state of interactions can be used to optimize the zone level energy consumption [19,20]. In order to integrate occupant workplace interactions into the control loop of the building systems, several strategies could be utilized: (1) Space utilization optimization. To improve energy efficiency of buildings systems through space utilization, occupants’ schedule of actions and interactions need to be integrated into a recommender system for assignment of space to the occupants and on-off adjustments on the building systems in spaces where services are not required. These savings could account up to 50% of total energy usages in under-utilized buildings. (2) Optimal space allocation. Occupants may have different habits in terms of overall space temperature and lighting levels. Given the distribution of the possible temperature level and lighting level as a function of building system operations, occupants can be guided to work in areas where the indoor environment fits their individual preferences and needs. Savings from space allocation is lower than space utilization due its milder constraints on space usages. (3) Personal thermal comfort driven building operations. Thermal comfort is defined as the state of mind which expresses satisfaction with thermal environment and it may change over time [22–26] and vary from person to person [27–29]. Optimal selection of HVAC temperature setpoints in office buildings can save energy up to 37% [21,30,31], although the savings differ based on the climate, building materials, and size. Given information on the HVAC energy consumption in a building, ambient sensing can be further extended to learn the occupant behavior that impacts building ventilation and heating/cooling (e.g., opening the windows, opening doors to adjacent spaces, preference for different air temperatures). This can then be used to control natural ventilation for conditioning and ventilation of the building [32–34] while considering available potential and limitations due to occupants’ behavior.

7. Discussion

In this paper, we introduce a framework for learning different occupant interactions based on ambient sensing. The type and quantity of occupant (social) interactions have impact on organizational productivity and can be used as a measured value when at the organizational level productivity is evaluated and suggestions for potential improvements are made. Due to the personal nature of measurement with wearable sensors, individual level of productivity can also be assessed. In the present study we were able to differentiate between working alone and working with other people, although we started with more comprehensive list of activities. In the follow-up work we intend to expand the list of activities and refine the occupant labeling process to increase prediction accuracy. It should be also noted that the dynamics of human behavior is a factor which is unmeasurable through large scale data collection but could be quantified and studied to better

---

**Table 3**

Features ranking for different binary work status labels.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Working alone</th>
<th>Working with other people</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Wall_sound</td>
<td>Wall_sound</td>
</tr>
<tr>
<td>2</td>
<td>Wear_sound</td>
<td>Wear_sound</td>
</tr>
<tr>
<td>3</td>
<td>Wall_CO2</td>
<td>Wall_CO2</td>
</tr>
<tr>
<td>4</td>
<td>Wall_pressure</td>
<td>Wall_pressure</td>
</tr>
<tr>
<td>5</td>
<td>Wall_relative_humidity</td>
<td>Wall_light</td>
</tr>
<tr>
<td>6</td>
<td>Wall_light</td>
<td>Wear_pressure</td>
</tr>
<tr>
<td>7</td>
<td>Wear_pressure</td>
<td>Wear_relative_humidity</td>
</tr>
<tr>
<td>8</td>
<td>Wall_temperature</td>
<td>Wall_relative_humidity</td>
</tr>
<tr>
<td>9</td>
<td>Wear_relative_humidity</td>
<td>Wear_temperature</td>
</tr>
<tr>
<td>12</td>
<td>Wear_light</td>
<td>Wear_light</td>
</tr>
<tr>
<td>13</td>
<td>Wear_temperature</td>
<td>Wear_temperature</td>
</tr>
<tr>
<td>14</td>
<td>Wear_CO2</td>
<td>Wear_CO2</td>
</tr>
</tbody>
</table>

* Corresponds to not useful features.

**Table 4**

Moving average results with different windows length.

<table>
<thead>
<tr>
<th>Window Length</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working alone (wearable and stationary)</td>
<td>0.8013</td>
<td>0.7962</td>
<td>0.764</td>
<td>0.7545</td>
</tr>
<tr>
<td>Working alone (wearable)</td>
<td>0.7219</td>
<td>0.7325</td>
<td>0.6646</td>
<td>0.6826</td>
</tr>
<tr>
<td>Working alone (stationary)</td>
<td>0.7815</td>
<td>0.7516</td>
<td>0.7453</td>
<td>0.7608</td>
</tr>
<tr>
<td>Working with other people (wearable and stationary)</td>
<td>0.8675</td>
<td>0.8333</td>
<td>0.821</td>
<td>0.8263</td>
</tr>
<tr>
<td>Working with other people (wearable)</td>
<td>0.7082</td>
<td>0.7756</td>
<td>0.7346</td>
<td>0.7303</td>
</tr>
<tr>
<td>Working with other people (stationary)</td>
<td>0.8411</td>
<td>0.8269</td>
<td>0.821</td>
<td>0.8383</td>
</tr>
</tbody>
</table>
understand which portion of the prediction error is associated with the dynamic behavior [35,36]. Needless to say, the reported high accuracy demonstrates that the compounded errors do not grow to significant values.

Learning and predicting patterns of space utilization are very important because maximizing space utilization directly saves energy that would otherwise be spent for environmental conditioning of unoccupied (and perhaps unneeded) indoor environments. A combination of stationary and wearable sensors can enable insight into space utilization and enable more effective allocation of available space [37]. In addition, wearable devices can house more sensors which can be used to learn personal comfort to better condition indoor spaces [25,34,38,39]. Occupant behavior in response to the utilization of new technologies in buildings has an impact on energy savings which needs further investigations [40–43]. There is also a trade-off between the complexity of sensing, learning and control paradigms and energy savings, which may play an important role for user adoption of these methods [44–46] which requires further investigation.

Since occupant interactions prediction is based on number of environmental indicators, the same information can be used for mapping spatial and personal environmental quality. The information quantifies the quality of the space and can be used for planning improvements. Wearable sensors can provide information about personal exposure during the whole day. This information can then be used to account for proper building ventilation needs and maximize health, wellbeing, and productivity. This type of information is missing from a number of past exposure studies that correlate health effects to exposure. The primary use of wearable sensor information will be in the realm of research until this type of information can be meaningfully used to inform ventilation rate adjustments based on the 24 h period exposure. Another useful application of sensing and learning interactions via wearable sensing technology is detection of contact and disease transmission networks [14,47,48]. This type of information can be useful in enabling engineering methods disease propagation mitigation (e.g., increasing ventilation supply where it is needed) or through non-engineering methods (suggesting to potential sources of infectious particles to leave the indoor environment or wear a mask). Finally, it should be noted that the proposed framework consists of different modules for modeling and modeling technique selection, feature selection, moving average sliding window length optimization, and label restructuring. Therefore, each module can be separately applied to problems that require a systematic approach to resolve an associated challenge. In our follow-up work we intend to focus on new applications of the described framework.

8. Limitations

Due to large scale data acquisition, this study was prone to errors on individuals’ actions. For example, we were not able to fully monitor each test subject’s usage of the wearable device, therefore there might be errors associated with some measurements due to occupants’ activity and behavior. It could be associated with the lower accuracies of predictions when using wearable sensing node. The data collection was carried out in mid-Atlantic and Southern states. Possible impacts of weather variations in multiple climates could not be integrated in the analysis. However, we should emphasize on the fact that our method focuses on the framework for learning the interactions, and therefore in any case that environmental conditions vary due to the occupants’ behavior, our framework could potentially work. As explained in the data collection section, we leave out some of choices in the user interface due to the fact enough positive responses for modeling were not reported. If we had proceeded to model those cases, we would have had overfit cases which are not desired. Another limitation in this study was the sampling rates. In order to enable long data collection, we collected data at averaged 5 min values. Even though we applied a moving window, and realize that for wearable devices, a longer averaged duration is optimal, we could not make any inference for shorter durations (smaller than 5 min average).

9. Conclusions

Real-time and continuous access to the state of interactions of building occupants enables design and control of smart building systems that optimizes both building performance and occupants’ connectivity. Such information would allow more complex building controllers to include various aspects of building occupants’ behaviors into the control logic. In this study, we proposed a framework for learning occupants’ interactions based on stationary and wearable sensing arrays. The framework also uses a recursive feature elimination method to select ambient sensing technologies which provide the optimal information gains and prevent overfitting. Accuracy, recall, and precision were used as performance measures for the interaction modeling, and the advantages of each one is discussed. Due to the fact that short term measurements of sensors might not correctly capture the dynamics of ambient conditions as it relates to the interactions, we include a component in the framework to generate different window length and select the optimal values. In cases where labels are mutually exclusive and could be coupled for a more in depth understanding of interactions in an environment, the framework utilizes a multi-label classification re-formulation. The framework was validated based on data collected in facilities managed by the US. General Services Administration in the mid-Atlantic and Southern states. 221 employees of federal agencies participated in this study and were randomly asked about their current task every hour. Sound (dB), CO2 (ppm), light intensity (lux), dry-bulb temperature (°C), relative humidity (RH%), pressure (mbar) measurements were collected from wearable devices worn around the neck by
the test subjects, and a network of stationary devices located in the test subjects’ working spaces. After cleaning the dataset, the framework was validated accordingly. The main variables for learning the interactions were: Wear sound, Wall sound, Wall CO2, Wall light. An overall accuracy of 86.72% for wearable and stationary devices, 81.25% for only-wearable, and 85.16% for only-stationary devices was observed for predicting the mixed multi-label classification via the Random Forests algorithm. Our findings suggest that office space activities that involve co-workers highly impact certain ambient conditions (e.g., sound levels and CO2), and therefore can be captured via simple stationary or wearable ambient sensing systems with minimal privacy concerns. In addition, these signatures are generalizable across large number of subjects. We plan to extend our research to capture patterns of personal variations in terms of their interactions. We plan to create profiles of interactions and cluster them into different group of workers.

Acknowledgement

This study was funded by the U.S. General Services Administration (GSA) under interagency agreement # GX0012829 with the U.S. Department of Energy and Lawrence Berkeley National Laboratory. GSA’s Wellbuilt for Wellbeing Group is a multidisciplinary research project team (GSA Contract # GS-00-H-14-AA-C-0094) consisting of the following members: Kevin Kampshoor, Judith Heerwagen and Brian Gilligan of GSA. Esther Sternberg, Perry Skeath, Casey Lindberg, and Matthias Mehlof the University of Arizona Institute on Place and Wellbeing. Bijan Najafi, Javad Razjouyan, Hyoki Lee, and Hung Nguyen of the Baylor College of Medicine Interdisciplinary Consortium on Advanced Motion Performance (ICAMP). Sudha Ram, Faizur Curim and Karthik Srinivasan of the University of Arizona INSITE Center for Business Intelligence and Analytics. Kelly Canada of LMI Inc. Priya Saha, Rebecca Goldfinger-Fein, Alicia Darbishire, and Mills Wallace of the Federal Occupational Health Service. Davida Herzel, Reuben Herzel, Melissa Lunden, Nicole Goebel, and Scott Andrews of Aclima Inc.

References