# OccuSpace: Towards a Robust Occupancy Prediction System for Activity Based Workplace

Mohammad Saiedur Rahaman*, Harsh Pare*, Jonathan Liono*, Flora D. Salim*, Yongli Ren*, Jeffrey Chan*, Shaw Kudo†, Tim Rawling†, Alex Sinickas†
*Computer Science and Information Technology, School of Science, RMIT University, Melbourne, Australia
†Arup
{saiedur.rahaman, harsh.pare, jonathan.liono, flora.salim, yongli.ren, jeffrey.chan}@rmit.edu.au
{shaw.kudo, tim.rawling, alex.sinickas}@arup.com

***Abstract*—Workplace occupancy detection is becoming increasingly important in large Activity Based Work (ABW) environments as it helps building and office management understand the utilisation and potential benefits of shared workplace. However, existing sensor-based technologies detect workstation occupancy in indoor spaces require extensive installation of hardware and maintenance incurring ongoing costs. Moreover, accuracy can depend on the specific seating styles of workers since the sensors are usually placed under the table or overhead. In this research, we provide a robust system called *OccuSpace* to predict occupancy of different atomic zones in large ABW environments. Unlike fixed sensors, *OccuSpace* uses statistical features engineered from Received Signal Strength Indicator (RSSI) of Bluetooth card beacons carried by workers while they are within the ABW environment. These features are used to train state-of-the-art machine learning algorithms for prediction task. We setup the experiment by deploying our system in a real-world open office environment. The experimental results show that *OccuSpace* is able to achieve a high accuracy for workplace occupancy prediction.**

***Index Terms*—Occupancy prediction, activity-based workplace.**

## I. Introduction

In an ABW, workers share desks and are encouraged to choose their seat anywhere from the available workstations. Over the past decade, the corporate world has seen an upward trend in ABW adoption [1], and several reasons have been attributed to this shift, including space and cost saving strategies to accommodate their workers. Also, the shared workplace has shown to bring greater brenifits including worker satisfaction [2], [3] and work flexibility [4]. Recent study highlights that the workers in shared workplace can exchange knowledge more effectively [5], [6] which can enrich their skills and enables them to be more productive [2]. However, it is very important to understand the utilisation of this shared workplace as it provides valuable insight for managers in strategic planning or accommodating the need for new spaces. One approach to achieve this goal is to predict occupancy of different atomic zones within a workplace [7].

The proliferation of smart building technologies has made it possible to sense, collect and analyse data related to buildings and its occupants. In occupancy analysis context, there are many types of sensors that can be installed under the desk or overhead to identify the desk level occupancy in the activity based work place. However, these technologies can be expensive and require extensive hardware installation. Also their maintenance can incur high levels of ongoing cost for the management. Further, the performance of these sensor based technologies are susceptible to noise and the occupancy detection accuracy relies too heavily on the workers' specific seating and movement behaviours. As a result, it is challenging to capture such behavioral data from the workers and provide occupancy prediction accordingly. In this paper, we collect relevant and valuable occupancy signals from a large ABW space by deploying a Bluetooth Low Energy (BLE) sensor network. Given a feature vector computed from the occupancy signals, we predict which atomic zone in the ABW space is occupied. We develop a system that utilises state-of-the-art machine learning techniques to learn and predict occupancy in different zones of the office space. The contributions of this paper are as follows:

- Deployment of a BLE sensor network to collect and process occupancy signals from different atomic zones in ABW environment.
- A robust system for noise-resistant indoor occupancy prediction through machine learning based modelling.

The paper is organized as follows. A summary of related works is given in Section 2. The prediction problem is defined in Section 3. We present our *OccuSpace* system for occupancy prediction in Section 4. The discussion of experiments and prediction results is given in Section 5. Finally, the paper concludes with a direction of future research in Section 6.

## II. Related Work

Prior researchers have explored the directions of occupancy pattern mining and prediction in work environments. A review of occupancy detection systems for building settings is presented in [8]. Additionally, researchers highlight the performance evaluation of chair sensors for occupancy detection in office building. In [9], an analytics approach to recognise the patterns of occupant presence is presented. The approach is based on cluster analysis that learns the rules of occupancy patterns by decision tree induction. The induced rules from decision tree are then used for occupancy schedule prediction.

* This paper is part of an ongoing research project supported by Arup and by use of the Nectar Research Cloud.

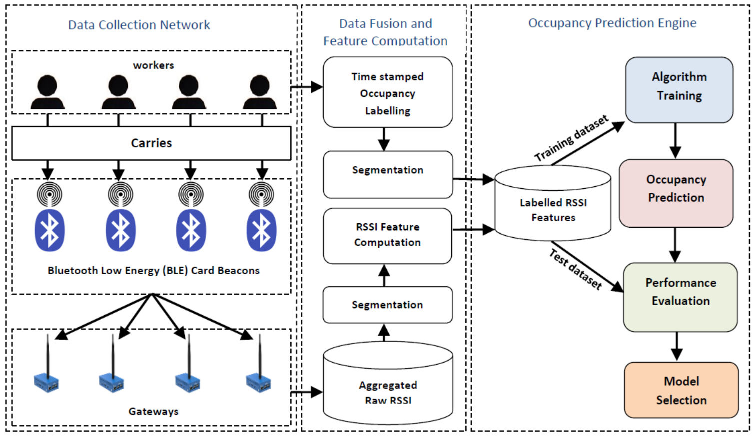


Fig. 1: OccuSpace System Overview.

Many sensor based techniques have also been proposed to understand the space utilisation in office space. In [10], the authors use $CO_2$ sensors to estimate the occupancy of indoor spaces. Another data-driven model based on decision tree and hidden Markov model (HMM) to predict occupancy using indoor environmental data is proposed in [11]. The need and design requirements of algorithms for occupancy prediction are first introduced in [12] and an extended version highlighting occupancy prediction in presence of multiple resolution is proposed in [13]. However, these devised techniques predict the overall occupancy of a relatively large indoor space and do not focus greatly on detailed movements of individuals in different atomic zones. To do this, further deployment of sensors is required. However, these sensor require costly installation and maintenance. Also, their occupancy prediction accuracy is highly dependant on specific seating styles of the occupants. Bluetooth sensors have become popular in recent times as they are cheap and able to collect data signals passively from indoor occupants [14]. However, traditional triangulation or trilateration techniques cannot be applied to Received Signal Strength Indicator (RSSI) data captured from BLE card beacons due to the omnidirectional nature of the signal coverage area. Instead, a grid-based fingerprinting technique can be used to calculate probable grid cell location of a device [15]. However, grid fingerprinting is usually conducted in controlled environments with equal grid sizes and presence of fewer people to avoid signal attenuation. Hence, this technique cannot be applied directly in large shared office environment. Other measures such as counting the total number of active MAC and IP addresses in the network [16] is not a reliable indicator of occupancy i.e. people may leave computers at their desks and usually use more than one device connected to the Wi-Fi access points.

Therefore, a system - *OccuSpace* is developed that can model noisy occupancy signals generated from workers' movement. More specifically, this paper uses Bluetooth RSSIs of moving BLE card beacons to train and compare machine learning techniques to predict occupancy in atomic ABW work zones of different sizes.

## III. Problem Definition

Let $Z = \{z_1, z_2, z_3, ..., z_n\}$ be the set of atomic zone labels that are created by dividing an activity based workplace into $n$ small blocks of different sizes. A zone is considered occupied when a worker carrying a Bluetooth beacon enters the zone and remains there for a period of time $\tau$. If $F^{rssi}$ is a feature vector computed from the Bluetooth beacon signal strength indicator over time $\tau$, we predict the occupied zone that corresponds to $F^{rssi}$. The goal is to induce a classifier $F^{rssi} \rightarrow Z$ based on the training samples described by the feature vector and the corresponding zone labels. We develop a system called *OccuSpace* to predict the occupied zone for a given feature vector.

## IV. The System Overview

The OccuSpace systems consists of three modules: i) data collection network, ii) data fusion and feature computation, iii) occupancy prediction engine. An overview of the OccuSpace system is illustrated in Figure 1 and the details of three OccuSpace modules are described below.

### A. Data Collection Network

This network collects data using two devices: Bluetooth gateways and Bluetooth Low Energy (BLE) card beacons from Kontakt.io as shown in Figures 2 (a) and (b). We deploy 4 Bluetooth gateways $(G_1, G_2, G_3, G_4,)$ at four different locations as shown in Figure 2 (c). The participant workers perform two tasks. First, the workers carry the assigned card beacons all the time while they are within the work environment. This allows the gateways to sense worker movements by capturing the variations in Received Signal Strength Indicator (RSSI). Second, the workers annotate their timestamped occupancy while they occupy a work station in one out of 28 atomic zones as shown in Figure 2 (c). The annotation is a zone label $z_i$ indicates that the $i^{th}$ atomic zone is occupied during any given time window. These annotations are used as ground truth for the occupancy prediction for different atomic zones. A brief description of BLE card beacons and gateways are given below:

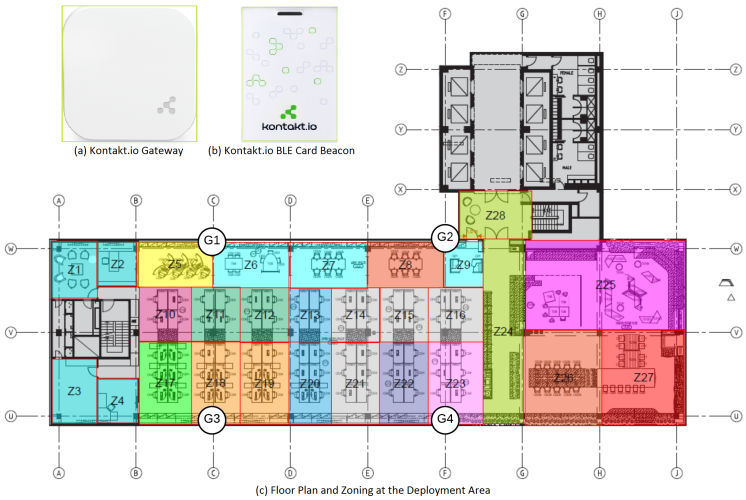


Fig. 2: (a) Kontakt.io Gateway (b) Kontakt.io BLE Card Beacon (c) Zoning by Dividing the Floor Plan of Our Deployment Site and Locations of 4 Gateways for Occupancy Prediction.

- BLE card beacons are transmitters that use Bluetooth Low Energy technology to transmit their unique identifier code to nearby receivers, in this case gateways. The transmitted unique beacon identifier (i.e. tracking id) can be used to detect the presence and movement of a person carrying a specific beacon. The BLE technology reduces power consumption thereby increasing beacon life span.
- A gateway device passively scans for BLE card beacons in its surrounding range. Upon detection of beacons, the gateway continuously collects beacon data including their unique identifiers, scan start and end timestamps, and RSSIs. The gateway uses RSSI to determine the proximity of the beacons in relation to the gateway. Note that the gateway(s) cannot use triangulation or trilateration techniques to localize beacons. Gateways send the collected beacon data to the cloud storage server approximately every 1.5 seconds.

### B. Data Fusion, Feature Computation and Mapping

In this stage, raw RSSI data from four gateways are aggregated based on the tracking ID of BLE card beacons. It should be noted that the collected RSSI data is continuous and spanned over a predefined time period. A short time period ($\tau$) is selected to proportionally segment the continuous RSSI readings into fixed windows. Subsequently, statistical features including maximum, minimum, mean, $1^{st}$ quartile, median, $3^{rd}$ quartile and variance are extracted from the RSSI data points in the temporal domain (bounded by $\tau$). Simultaneously, the timestamps of occupancy labels collected from the workers are also segmented based on $\tau$. Ultimately, the features and the timestamped occupancy labels of the workers are mapped to construct occupancy dataset with labelled RSSI features for occupancy prediction.

### C. Occupancy Prediction Engine

This module takes the occupancy labelled RSSI features dataset for occupancy prediction. We train a set of classifiers including J48, Decision Table (DT), $k$-Nearest Neighbor ($k$-NN), Support Vector Machine (SVM), Random Forest (RF), PART rules and Multi-layer Perceptron (MLP). The performance of each classifier is evaluated using the test dataset. We randomly divide the dataset into two parts: 60% for training the classifiers and 40% for testing the classifiers. The performance metrics used in the *OccuSpace* system is the predictive accuracy which measures the ratio of correctly predicted instances to the total number of instances evaluated. This measure can be used to find the top predictive model.

## V. Experiments and Results

We deploy *OccuSpace* in a large commercial building in Melbourne, Australia. We divide our deployment location into 28 atomic zones of interests. The zone boundaries are defined based on the seating arrangements of different teams during the study period. We have 4 participant workers who are assigned with individual Bluetooth card beacons. We set $\tau = 15$ and 30 seconds respectively for our experiment to examine if there are any variations in prediction performance for different occupancy time periods. We extracted 1820 and 1428 instances for $\tau = 15$ and 30 respectively. We found that the maximum predictive accuracy is produced by Random Forest (RF) which is just over 56% and 40% for $\tau = 15$ and 30 respectively.

We identified that this under performance is due to the fact that our dataset is unbalanced for different zone labels. By oversampling those instances and running the same experiment, we examine that the accuracies produce by some classifiers exceed 90%. This indicates that a good number of data samples are sufficient to predict the occupancy of atomic

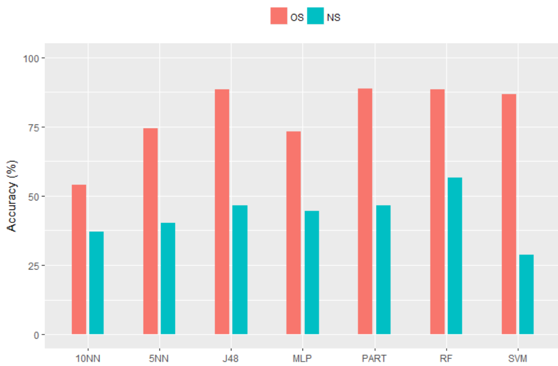


Fig. 3: Prediction Accuracy by Different Classifiers for $\tau = 15$ seconds ($OS$ = Over Sampling, $NS$ = No Sampling).

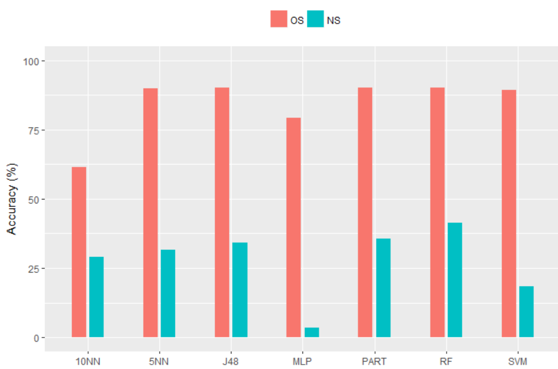


Fig. 4: Prediction Accuracy by Different Classifiers for $\tau = 30$ seconds ($OS$ = Over Sampling, $NS$ = No Sampling).

zones in an activity based office environment using machine learning techniques. Figures 3 and 4 show the accuracies of different classifiers using data over sampling (OS) and no sampling (NS) for $\tau = 15$ and 30 respectively. The systematic internal evaluation of classifiers is performed by averaging evaluation scores over $r$ iterations, where we use $r$ = 1000 as the parameter of the OccuSpace system.

## VI. Conclusions and Future Works

This paper presents a system for predicting occupancy in different atomic zones of different sizes in a real-world Activity Based Work (ABW) environment. A Bluetooth network was deployed to collect and process the occupancy signals from the occupant workers. A number of statistical features were computed from Bluetooth Received Signal Strength Indicator data which were used to train a suite of classifiers to predict the occupancy. We showed that a higher occupancy prediction accuracy can be achieved by using machine learning algorithms even within small time windows and with the presence of many workers in the environment. We also over-sampled the captured data to ensure balance across instances with different zone labels. The experiments showed that we can achieve more than 88.5% and 90% occupancy prediction accuracies for $\tau = 15$ and 30 seconds respectively when over sampling was applied. We observed RSSI signal drop for small period of time which may influence the computed features and thus cause prediction errors. By deploying more beacons can provide more granular data points. In future these issues can be addressed to increase the accuracy.

Currently we are devising techniques to compute zone-level occupancy numbers. Future work also includes spatio-temporal behavioral analysis of occupants in different ABW atomic zones. We also plan to investigate the occupancy patterns by integrating associated factors including lighting, noise, air quality and self-assessed productivity.

## References

[1] C. Wohlers and G. Hertel, "Choosing where to work at work towards a theoretical model of benefits and risks of activity-based flexible offices," *Ergonomics*, vol. 60, no. 4, pp. 467–486, 2017, pMID: 27167298.

[2] R. J. Cole, A. Bild, and A. Oliver, "The changing context of knowledge-based work: consequences for comfort, satisfaction and productivity," *Intelligent Buildings International*, vol. 4, no. 3, pp. 182–196, 2012.

[3] E. Sundstrom, R. E. Burt, and D. Kamp, "Privacy at work: Architectural correlates of job satisfaction and job performance," *The Academy of Management Journal*, vol. 23, no. 1, pp. 101–117, 1980.

[4] J. C. McElroy and P. C. Morrow, "Employee reactions to office redesign: A naturally occurring quasi-field experiment in a multi-generational setting," *Human Relations*, vol. 63, no. 5, pp. 609–636, 2010.

[5] S. J. Meijer EM, Frings-Dresen MH, "Effects of office innovation on office workers' health and performance," *Intelligent Buildings International*, vol. 4, no. 3, pp. 182–196, 2012.

[6] A. N. M., A. O. B., and J. K. A., "Understanding the physical environment of work and employee behavior: An affective events perspective," *Journal of Organizational Behavior*, vol. 35, no. 8, pp. 1169–1184.

[7] W. Wang, J. Chen, and X. Song, "Modeling and predicting occupancy profile in office space with a wi-fi probe-based dynamic markov time-window inference approach," *Building and Environment*, vol. 124, pp. 130 – 142, 2017.

[8] T. Labeodan, W. Zeiler, G. Boxem, and Y. Zhao, "Occupancy measurement in commercial office buildings for demand-driven control applicationsa survey and detection system evaluation," *Energy and Buildings*, vol. 93, pp. 303 – 314, 2015.

[9] X. Liang, T. Hong, and G. Q. Shen, "Occupancy data analytics and prediction: A case study," *Building and Environment*, vol. 102, pp. 179 – 192, 2016.

[10] I. B. A. Ang, F. D. Salim, and M. Hamilton, "Human occupancy recognition with multivariate ambient sensors," in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, March 2016, pp. 1–6.

[11] S. H. Ryu and H. J. Moon, "Development of an occupancy prediction model using indoor environmental data based on machine learning techniques," *Building and Environment*, vol. 107, pp. 1 – 9, 2016.

[12] F. C. Sangoboye and M. B. Kjærgaard, "Predicting occupancy presence in multiple resolutions for commercial buildings: Poster abstract," in *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*, ser. BuildSys '16. New York, NY, USA: ACM, 2016, pp. 229–230.

[13] F. C. Sangogboye and M. B. Kjærgaard, "Promt: predicting occupancy presence in multiple resolution with time-shift agnostic classification," *Computer Science - Research and Development*, vol. 33, no. 1, pp. 105–115, Feb 2018.

[14] W. Shao, F. D. Salim, T. Nguyen, and M. Youssef, "Who opened the room? device-free person identification using bluetooth signals in door access," in *Proc. of the IEEE CPSCom and IEEE SmartData*, 2017, pp. 68–75.

[15] H. J. P. Iglesias, V. Barral, and C. J. Escudero, "Indoor person localization system through rssi bluetooth fingerprinting," in *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, April 2012, pp. 40–43.

[16] R. Melfi, B. Rosenblum, B. Nordman, and K. Christensen, "Measuring building occupancy using existing network infrastructure," in *Proceedings of the 2011 International Green Computing Conference and Workshops*, ser. IGCC '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 1–8.