Indoor Occupancy Prediction using an IoT Platform

Alec Parise[1], Miguel A. Manso-Callejo[2], Hung Cao[1], Marco Mendonca[1], Harpreet Kohli[1], Monica Wachowicz[1]
[1]*People in Motion Lab, University of New Brunswick*
Fredericton, NB, Canada
[2]*Universidad Politécnica de Madrid*
Madrid, Spain
[1]{alec.parise, hcao3, marco.mendonca, harpreet.kohli, monicaw}@unb.ca, [2]m.manso@upm.es

***Abstract*—Current research in indoor sensor networks has pointed out an emerging interest in occupancy detection for Building Information Management (BIM) because buildings use 68% of Canadas energy in operation and contribute 17% of greenhouse gas (GHG) emissions. This research paper aims at developing a non-intrusive sensing method for predicting occupancy towards reducing building emission while also promoting a comfortable and productive working environment, while retaining the privacy of occupants. Towards this end, an IoT platform consisting of three main components: the edge computing environment, cloud based infrastructure, and network communication, together create a robust open source IoT architecture. The open source IoT architecture employs temperature, humidity, and pressure sensors for observing ambient environmental characteristics while combining PIR motion sensors, CO_2, and sound detectors. An occupancy detection model is then developed by applying Support Vector Machine (SVM) to predict occupancy patterns from the incoming IoT sensor data. This platform is a low-cost and highly scalable both in terms of the variety of on board sensors and portability of the sensor nodes, which makes it well suited for multiple applications related to occupancy and environmental monitoring.**

***Index Terms*—IoT Platform, Occupancy Prediction, SVM, data summarization, generic sensing**

I. Introduction

Indoor sensor network technologies for Internet of Things (IoT) have been a relevant research topic in academia for decades ranging from designing short range network infrastructures for data transmission and sensors APIs, for performing common sensor programming tasks regarding indoor localization, occupancy detection, ambient intelligence, and context-aware services [1]–[3]. The purpose of developing an indoor sensor network is to create an interoperable sensing environment able to establish communication between devices which have different data rates, latency, and network connectivity requirements. With advances in technologies such as the IoT, decision making AI infrastructures, and general connectivity of immersive technologies, indoor sensor networks are enabling devices to communicate any sort of "sensor data" through network carriers. In particular, the IoT aims to establish interoperability and interconnectivity among sensors to enhance human experience, especially those pertaining to physical contexts (e.g., home, office, and workshop) and the amenities contained within [4]. However, commercially available sensors are usually expensive and data access is restricted to the companies cloud and network infrastructures, which hamper their use due to a lack of interoperability. The proposed IoT platform for this research work has the ability to efficiently collect time series and event based raw data obtained from a variety of on board sensors measuring intrinsic environmental information for occupancy prediction. Current research in indoor sensor networks has pointed out an emerging interest in occupancy detection for Building Information Management (BIM), because buildings use 68% of Canada's energy in operation and contribute 17% of greenhouse gas (GHG) emissions [5]. As Canada pursues its COP21 agreement, there is a critical need not only to ensure that new buildings are designed to minimize energy consumption, but to address the performance of the much larger area of existing buildings. Several occupancy detection models have been proposed in the literature for lighting and heating control systems and optimization of Heating, Ventilation, and Air Conditioning systems (HVAC) that were based on a relatively small number of mathematical models such as Markov Chain, generic algorithms, fuzzy algorithms, and Artificial Neural Networks [6], [7]. However, previous occupancy detection models have focused on monitoring extensive simulation and scenario derived data, rather than real-world sensor data, mainly because they require costly platforms to monitor occupancy or they inaccurately monitor the occupancy due to irregular historical patterns.

This paper aims to propose an open source IoT platform that has the ability to efficiently collect time series and event based raw data obtained from a variety of on board sensors measuring intrinsic environmental information for occupancy prediction. The main scientific contribution of this research work is to avoid using simulated data and mathematical models, but instead explore data-driven models using real-world time-series and event data generated by employing an open source IoT platform to accurately predict when a room is going to be (un)occupied. The research premise is that data-driven models have the potential of improving our understanding of how indoor occupancy affects an environment and also how occupants behavior are impacted by the same environments. Towards this end, a Support Vector Machine (SVM) is used in the proposed platform. It is a well-known method and has been applied to many applications including classification of

This is a preprint. It will appear in *Proceeding of The 6th IEEE International Conference on Internet of Things: Systems, Management and Security,* Granada, Spain, October 22-25, 2019.

daily activities based on health sensor data [8], and activity recognition using event based trigger data [9]. SVM is a non-probabilistic binary classifier that relies on separating features via a hyper-plane and later classifies those features based on the distance of the maximum margins of the hyper-plane.

An indoor experiment was conducted using two general purpose sensing units that have yield diverse results, which were used to evaluate the proposed predictive model. The remainder of the paper is organized as follows: related work is presented in section II focusing on open source wireless sensing platforms and non intrusive sensing, section III present the proposed IoT platform, section IV describes the proposed algorithm for prediction, section V describes the implementation of the IoT architecture, section VI provides the preliminary results, and lastly in section VII the conclusions and future work.

II. Related Work

A. Open Source Wireless Sensing Platforms

The IoT relies on collecting information through a variety on sensing technologies including smart home devices, and wearable. Many of the off the shelf sensing platforms restrict cloud access and interoperability of other sensors. Very few attempts have been made to develop and deploy non-intrusive open source IoT sensor platforms for occupancy prediction. The literature in this section provides an overview of open source sensing platforms for energy monitoring systems, and relevant research for non-intrusive occupancy prediction using a SVM algorithm.

Ferdoush and Li [1] developed a low cost indoor Wireless Sensor Network (WSN) open source architecture utilizing a Raspberry Pi and an Arduino micro-controller. Temperature and humidity sensors were connected to an Arduino Uno which sent time-series data to a Raspberry Pi 3 acting as the base station. As a result of poor Wi-Fi connectivity this experiment used Xbee Wi-Fi modules (conforming to Zigbee networking protocols) as a means for Wi-Fi connectivity and for data transmission. When the data was received by the Base station it was stored locally in a MySQL relational database. This experiment proposes a pragmatic solution for personalizing an indoor wireless sensing network using a Raspberry Pi and an Arduino, however, rely on a limited sensor network for collecting real-world data.

Agarwal et al. [10] proposed energy management monitoring platform concentrating on occupancy to optimize the HVAC operating schedules. Multiple read switches and the Passive Infrared (PIR) sensors were both connected via WSN to interrupt enabled GPIO pins on a CC2530 micro-controller. To detect occupancy a rule associated algorithm was implemented based on possible scenarios. When a room was considered occupied the reed switch would have to detect that a door has been opened. This assumed that for a typical office building with an open door denotes the occupant is in the office or being somewhere nearby. Similarly, when people leave for an extended period (such as the end of the day) they typically close the door to their offices for security reasons. When a door closing event happens, there are two assumptions. Either the person closed the door and left (room unoccupied) or the person just closed the door and is still inside the room (room occupied). To disambiguate between these two cases the PIR sensor was used to determine if someone walked near the door. If the PIR sensor is triggered, then it is assumed there is still a person in the room. If the PIR sensor does not detect motion, then there is no occupant in the room. Based on these assumptions and the simulated experiment using EnergyPlus, they concluded that HVAC system operation costs were reduced by 10% to 15%. An issue with this assumption, however, is that a PIR motion sensor may not detect occupancy if the occupant is remaining stationary for long periods of time, therefore based on this experiments assumptions they will obtain erroneous data. Since PIR motion sensors only detect explicit movement past the sensor, an integration of other sensors such as sound and luminosity should complement the PIR sensor to negate false triggers.

B. Non-Intrusive Sensing

Abade et al. [11] proposed an open source framework for non-intrusive occupancy detection in an indoor environment based on four essential sensing units: temperature, CO_2, sound and luminosity. The IoT architecture for this experiment consisted of a Arduino Yun micro-controller which would send data every ten seconds to a Raspberry Pi. Then, aggregated means were taken of six values from each sensor and stored in a MySQL database. Logistic regression, SVM and Neural Networks were compared to individual sensor values to determine the accuracy of occupancy for each sensor. The results of this experiment concluded that logistic regression, followed by SVM yielded the most accurate results. The results of the temperature exceeded 89% for both SVM and logistic regression while CO_2 and noise sensors received accuracies of 6% and 1% for logistic regression and SVM. The luminosity was the best indicator of when the room was occupied as it obtained an occupancy detection accuracy of 95.60% for both machine learning algorithms. Due to occupancy behaviour some sensors would have far fewer events then others, for example luminosity would have far more events, especially if there is light pollution occurring from the windows. Therefore, having an accuracy of above 95% for occupancy detection may have been caused by external noise. Combining time series and event-based data allows us to determine the trends of the room and combing the sensor data will mitigate sensors which are triggered less often, which avoid hindering the accuracy for the occupancy prediction model. More importantly amalgamating the sensor data will create a more robust experiment, allowing the correlation of occupancy based on sensor data obtained by multiple sensor nodes to establish trigger sequence patterns.

Brennan et al. [12] suggested utilizing CO_2, temperature and humidity to estimate the occupancy by obtaining activities of daily living gathered by sensor nodes in a in a complex room inside a building. In total, six prototypes were placed in two laboratories. To filter and label the data the maximum and minimum distribution of observations from each sensor node was taken. For redundancy the authors, eliminated observations if the threshold of the CO_2 values did not exceed 5ppm. Samples were normalized on a per minute basis by collecting samples and taking the median of the observations for a 24-hour time window over a span of a week. Manual occupancy labelling was done based on the distribution of observations and peak measurements from CO_2 and humidity values. Brennan concluded the study by evaluating the correlation of CO_2, humidity, temperature, and through distribution of measurements was able to estimate occupancy within a room for smart systems.

Laput et al. [4] implemented a general-purpose sensing unit capable of gathering sensor data from an accelerometer, microphone, temperature, humidity, a 2.4 GHz Wi-Fi chip, and PIR motion sensor. This approach mainly relies on classifying microphone data for distinguishing the length of events and manually training an SVM model to classify specific simulated interactions such as: turning on a blender, microwave stove and faucet. The classification results of their experiment exceeded 85% accuracy; however only simulated events were used for training the SVM model. This research work has motivated our research on exploring SVM for predicting occupancy behavior, because it utilized a data driven event-based approach. By combing event-based data and time series it is possible to achieve a higher granularity given the random behaviors which occur over time. The implementation of this study was a major benefactor for how this experiment would tailor to different interactions within a room. Specifically, deploying a sound sensor capable of encapsulating object sound, and combining it with other on-board sensors such as motion to diversify the ability to classify room occupancy.

III. Proposed IoT Platform

The proposed architecture is illustrated in Figure 1. This section will describes each component that comprises the edge computing environment, cloud infrastructure, and means of communications between devices and the cloud.

A. Edge Computing Environment

Edge computing refers to the enabling technologies that allow computation to be performed at the edge of the network [13]. In other words, the edge grants the capability of performing simple data aggregations and analytical tasks to alleviate the network burden of transporting large amounts of data to the cloud or other data repository. Environmental sensors and gateways together create the edge computing environment. Once the data is sent to the gateways, it is then aggregated based on room location and sensor node ID and sent to the cloud.

B. Cloud Based Infrastructure

Cloud computing is a technology for computing resources provisions through the internet. It is a ubiquitous, convenient, network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) [14]. The services supported in cloud computing extend to data management/monitoring, storage, batch processing & analytics, scheduling, and auto scaling. The purpose of cloud computing and the relevance it has in this study pertains to the storage capabilities, using storage space and NoSQL database (i.e. MongoDB), and the ability to perform data cleaning and summarization tasks to train the predictive model (i.e. SVM) and eventually execute the predictive algorithm to predict user occupancy.

C. Network communication

In the proposed platform, the IoT devices communicate with the gateways through the short range, low power network communication technologies (e.g. Wi-Fi (IEEE 802.11), Bluetooth Low Energy 4.0). A direct wired connection to the internet was used for the gateway for two reasons: for time synchronization of the gateways to ensure there was no drift caused by the internal clock on each gateway, and for lower latency for transmitting the larger amounts of data from multiple sensors to the cloud for storage. To transmit the data from each device, a lightweight messaging protocol was used to publish and subscribe data packets from the sensors to the gateway and gateway to cloud.

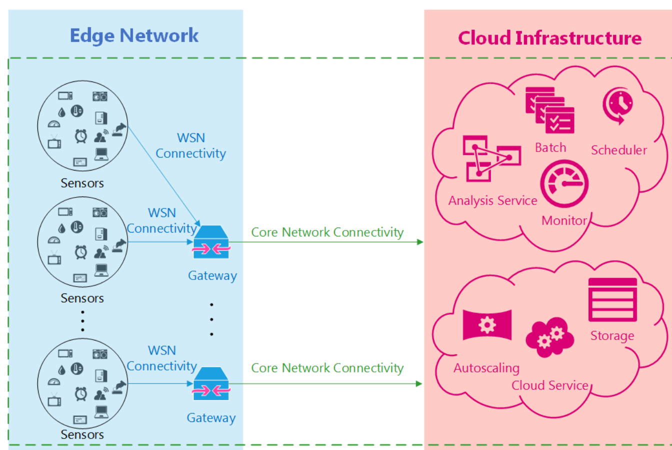

Fig. 1. The proposed IoT architecture for occupancy prediction.

IV. Occupancy Prediction Model

The proposed approach consists of deploying an open source IoT platform capable of streaming non-intrusive data to a private cloud service where the data is to be summarized and then analyzed. The data is collected from PIR motion sensors, temperature and CO_2 concentrations, humidity, pressure, sound detection and luminosity, all of which combined provided an efficient IoT platform for collecting raw sensor occupancy data and can be used in constructing an occupancy prediction model.

A. Data Summarization

Data summarization is a key data mining concept which involves techniques for finding a compact description of a dataset. It has been proven to be a useful and an effective technique supporting data analysis of large amounts of IoT data. Building predictive models from IoT data is time consuming, and summarization is an important step to expedite analytical tasks by reducing the size of processed data.

Given a set of n categorical features $F = \{F_1, F_2, ..., F_n\}$ and an associated weight vector W such that each $W_i \in W$ represents the weight of the feature $F_i \in F$. A set of transactions T, such that $|T| = m$, is defined using these features, and each $T_i \in T$ has a specific value for each of the n features. Formally, a summary of a set of transactions can be defined as follows: A summary S of a set of transactions T, is a set of individual summaries $\{S_1, S_2, ..., S_k\}$ such that *(i) each S_j represents a subset of T* and *(ii) every transaction $T_i \in T$ is represented by at least one $S_j \in S$.*

B. SVM Prediction Model

SVM is a supervised learning algorithm which relies on separating objects via hyper-plane and selecting the appropriate event by successively shrinking or reducing a dataset to determine which objects are closest to the hyper-plane [15]. In this algorithm, each data item (i.e. measurement) is plotted as a point in n-dimensional space (where n is the number of features in the dataset) with the value of each feature being the value of a particular coordinate. Support Vectors are simply the coordinates of individual observation and SVM is the point which best segregates by the two classes via hyper-plane. Classification occurs by inputting a training sample set in the case of this experiment 80% of the original dataset was used. Then the algorithm can be parameterized to find the best fit hyper-plane that differentiates the two binary classes.

V. The IoT Architecture Implementation

In this section the implemented architecture for the IoT platform will be discussed along the components which comprise the edge computing environment, the cloud based infrastructure and the network connectivity. The overall architecture is illustrated in Figure 2.

A. Implementation at the Edge

The edge computing environment consist of two main elements: the sensors and the gateways. The sensor nodes were designed to represented two unique environments which encompass different types of occupancy patterns; entry/exit and prolonged occupancy. The nodes contained six sensors consisting of temperature (BMP280), humidity (HC1080), luminosity (SI1145), motion (HC-SR501), sound (KY-038) and CO_2/TVOC (SGP30) sensors. The sensors are connected to a NodeMCU micro-controller that sends via WiFi all observations periodically (i.e. in 10s intervals) or when events occur due to an occupant triggering the sensors. The data is sent using publish/subscribe (details in section V-C), to an MQTT broker through the Raspberry Pi 3 B+ acting as the gateway for one or more NodeMCUs. At the gateway, data is consolidated based on the sensors ID and room location. Each tuple is then designated a timestamp for when the data arrives at the gateway and is sent to the cloud.

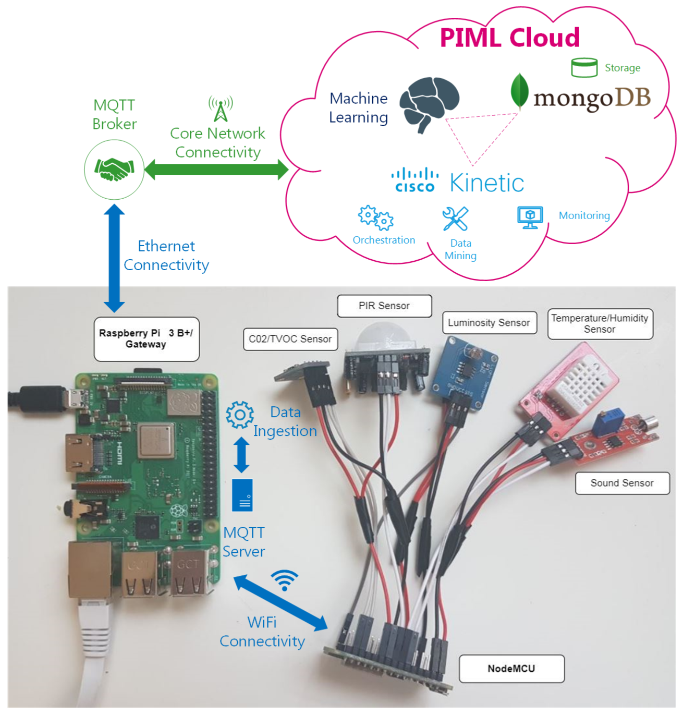

Fig. 2. Implemented IoT architecture for occupancy prediction.

B. Implementation in the Cloud

It is complicated to handle massive amount of data tuples coming to the cloud from different distributed gateways. Therefore, Cisco Kinetic has been utilized for the incoming IoT data streams management in the cloud resource. This data streams management platform includes three key modules: Gateway Management Module (GMM), Edge & Fog Processing Module (EFM), and Data Control Module (DCM). Cisco Kinetic allows for *(i) extraction of data from disparate sources ("things"); (ii) compute data anywhere from edge to destination and provide processing where it is needed; and (iii) move data programmatically to get the right data to the right applications at the right time*. It has also been used to implement the data summarization tasks needed for monitoring the IoT data streams before using them in the occupancy prediction model (Figure 3).

Fig. 3. Cisco Kinetic Dataflow Layout.

C. Network Communication

Communication between the NodeMCU and Gateway was configured using Wi-Fi connectivity, operating in 802.11 protocol with Wi-Fi frequency band of 2.4 GHz. A bandwidth of 2.4 GHz was preferred due to its longer range which allowed for a larger coverage of the gateways, which is an important consideration for scalability of the experiment. A publish/subscribe, an Internet of Things connectivity protocol, was used to send and retrieve tuples from an HiveMQ broker. In other words the NodeMCU client would publish the data to the MQTT broker and the gateway would subscribe to retrieve the data. Direct wired connection to the internet was used for the gateway for time synchronization of the Raspberry Pi, and for lower latency when transmitting larger amounts of data. To send the data tuples from the gateway to the cloud, the publish/subscribe protocol was utilized once again to send and retrieve data from the HiveMQ broker.

D. The Design of Experiment

The test selected for this experiment was conducted in a windowless classroom on the Geodesy and Geomatics Engineering department floor, which included a lecture podium, projector and twelve desks each seating two people. The experiment spanned over a two-week duration and was conducted during the winter semester to encompass group activities at scheduled instances to assess the accuracy of the labeled class. The first sensor node was placed at the doorway and the second at the front of the room beside the lecture podium, seen in Figure 4. The placement of the first sensor was designed to encapsulate entry and exit events while the other sensor node was able to detect when movement occurred by the lecture podium, indicating when a lecture was in session. Since there were only two prototype sensors, it was important to place them in areas of high activity, while also maximizing coverage, and avoiding sparse Wi-Fi connectivity. Furthermore to ensure that the gateway had consistent connection to the two sensor nodes, a preliminary experiment was performed to determine how far the gateway could be placed before the connectivity became slightly disrupted.

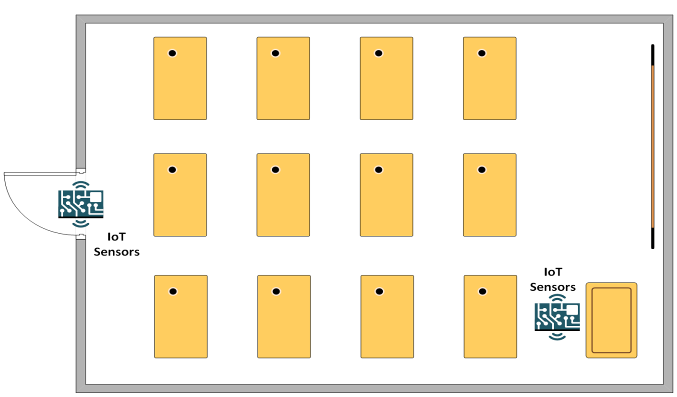

Fig. 4. The indoor room configuration and location of the NodeMCUs

VI. Preliminary Results

The experiment was conducted over a span of 13 days, during the experiment data summarization tasks were performed every time a tuple arrived in the cloud. The summarization tasks included, a computations of the average values of the classroom observations, the standard deviations, and the computed output of maximum temperature being observed in the classroom. This was important to monitor the the classrooms typical ambient values to differentiate when the room was occupied or unoccupied.

The initial triggers were measured by the sensor node placed at the doorway which consisted of PIR, sound and CO_2 sensors, specifically designed to capture when people have entered the classroom. To facilitate the sensor node placed the door was another sensor node placed at the front of the classroom; consisting of six sensors cited previously (temperature, humidity, sound, luminosity, motion, and CO_2) and was designed to measure when a lecture was in progress. Data was merged in the gateway and stored in the cloud to later be used for the SVM predictive model. To build the model 80% of the data set was selected for training while the remainder was used the testing data set. The results of the predictive model experiment can be seen in Table I.

TABLE I
THE EVALUATION OF OUR SVM PREDICTIVE MODEL.

	Precision	Recall	F1 Score	Support
Free	0.99	0.97	0.98	40353
Occupied	0.87	0.95	0.91	9186
Micro Average	0.96	0.96	0.96	49539
Macro Average	0.93	0.96	0.94	49539
Weighted Average	0.97	0.96	0.96	49539

The overall predictive accuracy was 96%, but to truly understand the performance of the algorithm, we must look at the measurements involved throughout the process; these include precision, recall and F-1 score. The model was considerably accurate when predicting when the class room was free. The F1-score was the focal measurement because it is a representation of both the recall and precision of the SVM model. To further elaborate on the performance of the algorithm, a confusion matrix was implemented for understanding how many misclassified observations had occurred when testing the predictive model, seen in Table II. The matrix can be divided into four categories, True Positives, False Negative, False Positive and True Negatives. The properly predicted observation fall into the True Positive and True Negative categories, and based on the results show that 2% of the Free statuses were misclassifed and 5% for the Occupied predicted statuses. The results displayed a higher achieved accuracy that the literature presented, and as a consequence the predictive model obtained an overall accuracy of 96% and has the potential to be used for schedule optimization towards reducing building emissions and to facilitate a more comfortable environment for occupants while retaining user privacy.

TABLE II
CONFUSION MATRIX RESULTS.

	Free	Occupied
Free	TP = 39002	FN = 1351
Occupied	FP = 474	TN = 8712

VII. Conclusion and Future Work

Indoor wireless sensor networks are key components for understanding how people use facilities within a building. Research in todays indoor wireless sensing networks rely on collecting intrusive information for validating occupancy or as a means for monitoring energy consumption within the building using mathematical or simulated models. The purpose of this study was to deploy an interoperable non-intrusive sensing network capable of encapsulating patterns of occupancy towards behavioural pattern recognition and prediction for building information management, emergency protocols and general promotion of the facilities within a building. Two sensor nodes were programmed and deployed, each sending data in near-real time to a Raspberry Pi (gateway), and eventually stored in the cloud for analytics. The data was consolidated and analyzed using a popular classification and predictive machine learning algorithm (SVM). The results presented in the previous section exceeded expectations by achieving a higher accuracy then any of the presented literature, while retaining the security of occupants. Since the occupancy prediction model was created as a binary classifier, SVM proved to be an efficient algorithm both computationally and in terms of performance, yielding a predictive accuracy of 96%. This experiment provided useful insights for threshold tuning of sensors and parameter adjustments to optimize the SVM algorithm.

For future research, adjustments for the CO_2 thresholds will be made before scaling the sensor deployment throughout the engineering complex. Another consideration for scaling up the deployment of the IoT indoor sensing network will be to geographically determine the best locations for the sensors. Each environment is unique and will yield different results and by understanding the most optimal locations the predictive model will be less susceptible to false predictions. Apart from adjusting the thresholds and sensor deployment, creating a more concise occupancy status label for training the SVM model will also be a consideration for future research. The current label occasionally mis-classifies occupancy statuses when individuals are occupying a room, therefore optimizing the label will yield a more accurate occupancy prediction. Furthermore, integrating regression techniques into the analytical pipeline will associate time-stamps with the predicted occupancy patterns, which will be fundamental for future research. The research presented is a novel approach for occupancy prediction using non-intrusive information and the proposed IoT indoor wireless networks is highly scalable and can be deployed on multiple floor levels and buildings due to the interoperability of sensors nodes, gateway and cloud. In the future, other technologies will be integrated into the architecture such as Cisco Edge and Fog nodes. These technologies will be a primary component from streaming analytics and pave the way for real-time sensor data analytics. The applicability of this research is not only contained to a room, but can be used to analyze patterns of occupancy at the floor level, building, and campus. This will contribute to optimizing operation times of HVAC systems for reducing CO_2 emissions, promote a comfortable environment for occupants, and retain user transparency of privacy.

Acknowledgments

This work was supported by the NSERC/Cisco Industrial Research Chair, Grant IRCPJ 488403-1. We would like to thank Compute Canada for providing us with the cloud resources.

References

[1] S. Ferdoush and X. Li, "Wireless sensor network system design using raspberry pi and arduino for environmental monitoring applications," *Procedia Computer Science*, vol. 34, pp. 103–110, 2014.

[2] G. Anastasi, R. Bandelloni, M. Conti, F. Delmastro, E. Gregori, and G. Mainetto, "Experimenting an indoor bluetooth-based positioning service," in *23rd International Conference on Distributed Computing Systems Workshops, 2003. Proceedings.* IEEE, 2003, pp. 480–483.

[3] M. Jin, S. Liu, S. Schiavon, and C. Spanos, "Automated mobile sensing: Towards high-granularity agile indoor environmental quality monitoring," *Building and Environment*, vol. 127, pp. 268–276, 2018.

[4] G. Laput, Y. Zhang, and C. Harrison, "Synthetic sensors: Towards general-purpose sensing," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* ACM, 2017, pp. 3986–3999.

[5] D. Schulte, "Better buildings for a low-carbon future," https://www.ourcommons.ca/Content/Committee/421/ENVI/Reports/RP9989842/envirp17/envirp17-e.pdf, 2018, [Online; Accessed on 2019-09-04].

[6] 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.

[7] W. Wang, J. Chen, and T. Hong, "Occupancy prediction through machine learning and data fusion of environmental sensing and wi-fi sensing in buildings," *Automation in Construction*, vol. 94, pp. 233–243, 2018.

[8] A. Fleury, M. Vacher, and N. Noury, "Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results," *IEEE transactions on information technology in biomedicine*, vol. 14, no. 2, pp. 274–283, 2009.

[9] N. C. Krishnan and D. J. Cook, "Activity recognition on streaming sensor data," *Pervasive and mobile computing*, vol. 10, pp. 138–154, 2014.

[10] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, and T. Weng, "Occupancy-driven energy management for smart building automation," in *Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building.* ACM, 2010, pp. 1–6.

[11] B. Abade, D. Perez Abreu, and M. Curado, "A non-intrusive approach for indoor occupancy detection in smart environments," *Sensors*, vol. 18, no. 11, p. 3953, 2018.

[12] C. Brennan, G. W. Taylor, and P. Spachos, "Distributed sensor network for indirect occupancy measurement in smart buildings," in *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC).* IEEE, 2018, pp. 1290–1295.

[13] H. Cao, M. Wachowicz, C. Renso, and E. Carlini, "Analytics everywhere: generating insights from the internet of things," *IEEE Access*, 2019.

[14] G. Zhang and M. Ravishankar, "Exploring vendor capabilities in the cloud environment: A case study of alibaba cloud computing," *Information & Management*, vol. 56, no. 3, pp. 343–355, 2019.

[15] T. Joachims, “Making large-scale svm learning practical,” Technical report, SFB 475: Komplexitätsreduktion in Multivariaten , Tech. Rep., 1998.