Demonstrating the potential of indoor positioning for monitoring building occupancy through ecologically valid trials

Md. Shadab Mashuk ^{1,*}, James Pinchin ², Peer-Olaf Siebers ³ and Terry Moore ⁴

^{1,4} Nottingham Geospatial Institute; <u>shadab.mashuk@nottingham.ac.uk</u>, <u>terry.moore@nottingham.ac.uk</u>
 ² Department of Architecture and Built Environment; <u>james.pinchin@nottingham.ac.uk</u>
 ³ School of Computer Science; <u>peer-olaf.siebers@nottingham.ac.uk</u>
 * Correspondence: <u>shadh06@gmail.com</u>

Abstract: Assessing building performance related to energy consumption in post-design-occupancy stage requires knowledge of building occupancy pattern. This occupancy data can potentially be collected from trials and used to improve the prediction capability of building performance models. Due to the limitations of passive sensors in detecting an individual's occupancy throughout the building, indoor positioning can provide a viable alternative. Previous work on using indoor positioning techniques for detecting building occupancy mainly focused on passive monitoring through Wi-Fi or BLE proximity sensing by estimating the number of occupancy monitoring through active tracking at an individual level using a smartphone-based multi-floor indoor positioning system. The paper discusses the design of a novel occupancy detection trial setup, mimicking real-world office occupancy and discusses the outcome of the ecologically valid trials using the developed positioning system. In total 50 occupancy trials were carried out by around 22 participants comprising of a variety of routes within the building. The trial results are presented to demonstrate the level of accuracy achievable against a specific set of the performance metric necessary for building occupancy detection and modelling.

Keywords

Building Occupancy; Building Performance; Indoor Positioning; Smartphone; Multisensor Fusion

1. Introduction

The urbanisation of the world is occurring at an unprecedented scale and is being affected by various local and global factors. Quality of life, environment, and policy related to the regulation of energy supply will be severely affected if sustainable urban growth is not ensured. Building energy models (BEM) has historically been the mode of understanding and analysis of energy performances of municipalities and neighbourhoods for building designers and energy policymakers [1]. This increasing demand for more energy-efficient buildings has created a challenge for the construction industry to

ensure the building energy performance predicted during the design stage sustains once the building is in use. Unfortunately, current statistics provide evidence to suggest that buildings are not performing as well as expected [2] and that the predicted performance of a building at the design stage is often not matched when the building is in use which is known as the "performance gap". The reason for this performance gap is due to various discrepancies related to design and model assumptions, management of power sources, occupancy behaviour and quality of building materials. Initiatives such as Post Occupancy Evaluation (POE) and Post-occupancy Review of Buildings and their Engineering (PROBE) aims to illustrate the extent of this so-called "performance gap" [3]. Typically POE takes place after twelve months of building occupancy, and in all the cases prediction tends to be unrealistically low while actual energy performance is usually unnecessarily high. Many of the causal factors for relate to the use of unrealistic input parameters in the model regarding occupancy behaviour, energy facilities management and usage patterns and further associated with the lack of feedback to designers once a building has been constructed and occupied. Results from a case study using the PROBE suggested the measured electricity demands are approximately 60–70% higher than predicted in schools and offices, and over 85% higher than predicted in university campuses [2]. As discussed in [4], [5] the energy-use behaviour in commercial buildings depends on a lot of factors such as occupancy pattern, occupant actions and interactions with building infrastructures and appliances. Through occupancy monitoring, we can identify occupancy numbers and patterns which acts as a key input for building simulation tools such as Energy Plus, ESPR, DeST and TRNSYS, producing energy consumption forecast of heating, cooling and lighting. The forecasting accuracy depends on the accuracy of occupancy dynamics recorded from monitoring. It also acts as input to model complex stochastic behaviours such as light switching, window opening and device usage, producing probabilities of performing an action within the building simulation environment. Occupancy patterns also varies because of office demographics and occupant entry and exit time, short and long breaks and out of hour stay contributing significantly to variation of energy-use behaviour. Thus it is important to associate occupancy with energy usage both at an aggregate and individual level. Existing literature shows occupancy monitoring encompasses a wide range of detection technologies such as camera, CO_2 sensors, humidity sensors, infrared (IR) sensors, light sensors, motion sensors, radio frequency identification (RFID), sound sensors, temperature sensors and Wi-Fi infrastructures [4]. Although aggregate occupancy can be easily determined using the above technologies, it is quite complicated to detect occupancy at an individual level which is necessary to identify variation from a demographics perspective as well as provide occupancy information at a greater detail providing both zone level and room level accuracy including transitions between rooms, corridors and stairs. As such the research goal is to analyse the performance and understand the limitations of the proposed IP system for occupancy monitoring at an individual level when carried out in a real-world setting.

The paper demonstrates a novel application of indoor positioning (IP) solution capable of detecting building occupancy at an individual level continuously with sufficient accuracy and detail (room level

and zone level) over a long period. Thus providing post-processed location-based services using the collected data related to building performance modelling and simulation. It is important to note that this paper is an extension of our previous study [6] where a multi-floor IP system was implemented and validated hence we will not repeat the technicalities of the positioning system design any further.

We propose a novel occupancy trial design and use the developed IP system from our previous research [6] to carry out "extensive" multi-user occupancy trial to understand its performance in an unbiased and objective manner in a real-world setting to ensure ecological validity. The occupancy trial takes the form of a spatially distributed Lego model-building game to imitate real-world occupancy in office buildings instead of providing the participants with a list of places to visit. This helps to create the environment for the ecological validity of the trials by ensuring participants are more focused on the Lego game and resource collection while moving freely and naturally and not being too aware of the positioning task and data collection process through the smartphone. The insights gained from the trials is a major contribution of this paper. The accuracy of the detected occupancy from the trials is discussed and compared with existing occupancy detection technologies. The next two sections will discuss related work and methodology adopted. We then provide a thorough analysis of the ecologically valid trial results and performance followed by a discussion and conclusion.

2. Related Work

There has been a lot of work on building occupancy detection and modelling; adopting different methodology such as probabilistic modelling approaches [7]–[9], use of environmental sensors [10], [11], localisation sensors [12], [13] and agent-based modelling approach [14]. Development of sophisticated occupancy and occupant model requires collection and analysis of accurate occupancy data that helps to identify occupancy patterns in detail. An early example of occupancy detection using passive infrared (PIR) sensors can be seen in [15], [16] where the authors carried out occupancy detection experiments in singly occupied office rooms and used the data to develop occupant presence model. Ambient sensors, as well as CO₂ and related sensors, can also be used to capture the occupancy status of buildings. Environmental signatures such as light, humidity, CO₂ and temperature are important indicators of occupancy in buildings [4], [10], [11], [17]. Statistical modelling and machine learning algorithms such as decision tree, artificial neural network (ANN), support vector machine (SVM), and random forest can be used to predict occupancy level, learn patterns of occupancy and develop occupancy models [10], [11], [18]. Combination of environmental sensors and camera technology is also quite popular for detecting building occupancy [19]. Use of localisation sensors and wireless communication technology can be seen to explore the area of occupancy detection. Wi-Fi technology, widely available at home and in offices seems to be a feasible choice to determine occupancy. Using statistical modelling and data mining techniques we can estimate the number of occupants based on the location of the detected AP's and connected devices of individuals as seen in [12], [13]. Bluetooth Low Energy (BLE) with its customizable range and low cost can be a very effective solution for occupancy detection. Individuals with smartphones can be detected in each room and using the configured range the location of the individual can be identified providing detailed occupancy information as seen in [20]. Although the authors did not carry out any extensive occupancy trials and were limited to testing on a small scale within a single room only. Probabilistic approaches in developing occupancy models such as graphical models and Hidden Markov models (HMM) are discussed in [9], [21], [22]. Use of tactile sensors in chairs such as commercially available embedded low-cost microswitches can be found in [23]. A detailed review paper of occupant detection systems, modelling approaches and their performance evaluation can also be found in [24] where the advantages and limitations of these modelling methods are also compared and analyzed, as well as appropriate recommendations are made for future studies. One of the important aspects of classifying existing occupancy monitoring technologies is the spatial and temporal resolution of detection. As seen in [4], [17] the authors highlighted the need to identify the state of occupants, count, activity and tracking and also whether the temporal resolution is in seconds, minutes, hours or days and if the detection is limited to rooms, floors or the entire building. In both the papers the authors provide a detailed review of existing occupancy modelling techniques such as analytical approach using algorithms such as steadystate algorithm (ST-ST) or dynamic algorithm (dC) in combination with CO₂ sensor data, knowledgebased approach represented by specialist rules based on the sensor data collected as well as data driven approach using a wide range of machine learning and probabilistic algorithms making use of data collected from Wi-Fi, BLE, PIR and CO₂ sensors, smart meters etc. They also discuss simulation of occupant energy consumption behavior using the occupancy monitoring data as input in agent based model (ABM) or multi agents systems (MAS) and other methods combining machine learning and probabilistic techniques coupled together.

Some of the shortcomings of the occupancy detection methods mentioned above are that the trials can only detect the general occupancy of rooms and corridors. It becomes very difficult to quantify the actual number of occupants directly from the sensor data although model predictions recreating patterns of occupancy distributions vary between 60 - 90 % as seen in [17]. The PIR sensor's detection is limited to individual office rooms only thus it will not be possible to know the whereabouts of the individual once they exit their respective office room and collecting data sets to train probabilistic and machine learning models are time-consuming and can be computationally expensive. Deployment cost, time and maintenance of the system is a major factor when carrying out trials and data collection covering entire buildings over long periods with a sizable population similar to any commercial office environment. Also, occupant demographic information and detailed transition during occupancy are difficult and to some extent impossible to monitor and unlikely to be cost-effective using PIR, RFID and environmental sensors. Thus we are missing the varying patterns of occupancy exhibited by different individuals not only in their respective offices but also during the transition from one room to another through stairs and corridors. The modelling also suffers from challenges due to oversimplification of the input parameters and may require significant calibration. As a result, they cannot be generalised for other building types and context thus missing out the ecological validity necessary to generalize the outcome for other similar buildings.

To develop more realistic and accurate occupant models for building performance simulation (BPS), IP technology can play a very constructive role in monitoring occupancy in large commercial buildings with greater detail and develop movement profiles at an individual as well as aggregate level with sufficient accuracy. Wi-Fi and BLE have large overlapping area coverage, widely available and cost-effective compared to other technologies and likely to provide transition information within zone level accuracy if the range is carefully configured and deployed. As such, good for proximity-based monitoring but adding pedestrian dead reckoning (PDR) capability will allow continuous active tracking of occupants recording detail information of transitions between each room and corridors throughout the entire building. Thus creating movement profiles at an individual level and also identify varying patterns of occupant behaviour based on occupant demographics. These movement profiles of individuals can be used to improve the BPS model by adding more detail to the simulation of energy usage prediction and also act as input for other behavioural models such as appliance usage, window opening, light switching, and shading device usage [25]–[28]. Thus we must try to evaluate the feasibility of carrying out occupancy monitoring through active tracking using the IP system with PDR capability in a real-world environment.

In the next section we discuss in detail the methodology of the occupancy trial setup in the form of a Lego building task; spatially distributed and designed to mimic real-world occupancy scenario in an office building involving non-expert participants.

3. Methodology

3.1. Occupancy Trial Test Bed Setup

Building occupancy implies being physically present inside the building and performing day to day activities and tasks with movement between rooms and corridors from time to time. To carry out the occupancy monitoring trials we make use of the IP system developed in our earlier research in [6] with the deployed BLE beacons and Wi-Fi AP's covering Floor A and Floor B of the Nottingham Geospatial Building (NGB) as seen in Figure 1 and 2. The trials aim to detect the sequence of the participant's movement and transition through rooms and corridors during the task and were carried out covering both the floors. The general flow of the trial can be understood from Figure 3 starting from initialisation to occupancy computation. Data collection during the trial was done through a data logger using android LG Nexus 5 smartphone by all the participants. The data logger continuously captures Wi-Fi, BLE, accelerometer and gyro data as defined in our previous work [6] and post-processed in the IP Location Engine as seen in Figure 4 to compute the sequence of occupancy.

Each of the participants was asked to complete a list of spatially distributed task without any guidance from the researcher. They had access to all the rooms, corridors and stairways of the two floors when completing the task and they could move freely. One of the major constraints was participants were

asked not to run or jog as the motion detection algorithm implemented could only detect walk-like motion; slow, medium and fast-paced in this research. Also since the orientation of the phone continuously changes during walking, we would need to recalibrate the heading continuously by detecting the tilt which is a challenge. We have decided to fix the phone orientation in the pocket as a limitation in this research and just computed heading change rather than absolute heading. The general idea was to ensure real-life building occupancy as much as possible without distracting the participants during the trial and evaluate the accuracy of occupancy detection within room level and zone level accuracy.



Figure 1. Floor A showing the BLE beacons in green and Wi-Fi AP's in red

Figure 2. Floor B showing the BLE beacons in green and Wi-Fi AP's in red



Figure 3. Block diagram of the occupancy trial flow



Figure 4. Block diagram of the Positioning System Location Engine

3.1.1. Trial Objectives

Determining an IP solution is complicated because of varying context of their application and userdependent performance requirement. Different technology, including hybrid systems, may combine multiple sensors with an appropriate measurement principle and algorithms for determining the positioning solution [29], [30]. IP application areas also have a lot of overlap when it comes to coverage and accuracy and may come with inherent constraints related to technological capability, deployment issues and options for modifications such as added functionality. Also, some of the most critical performance-defining parameters that are generally looked for in an IP solution is not limited to accuracy, area coverage, latency, availability, robustness, data accessibility and security but most importantly cost. All these parameters need to be addressed when choosing a positioning technology and matched with the user requirements. To make things more complicated, these parameters are not readily determinable beforehand since they are in turn dependent on environmental factors, assumptions and the application domain. Thus it is challenging to determine the best possible solution objectively and generally requires an iterative trial and error process for any real-world application as seen in [29], [31], [32]. Likewise setting up the trial environment correctly to test the IP solution is also a challenge since it dictates the efficacy of the trial outcome. In typical IP trials, experiments are carried out in a controlled environment, unlike real-world settings but to understand the quality of building occupancy trial performances this may not be relevant. So far there has been no previous work on occupancy

monitoring through continuous active tracking over a long period. As such Ecological validity is the best way to ensure test performance predicts behaviours in a real-world setting. Ecological validity implies the observation seen in experimental or laboratory setup can be generalised in other similar natural settings [33], [34], in our case the performance observed from the trials in our test building site should be what we can expect in any other building with any other participants in a similar setup. Thus to ensure ecological validity we have to make sure the environmental setup and context of the trial is similar or as much closer as possible to real-world occupancy. Also, the nature of the task, behaviour and response of the participants should be unbiased and similar to an office environment [35]. Since active tracking will take place over a longer period it is important to ensure the positioning system is robust and precise when computing occupancy of participants from the recorded positioning data. It also needs to take into account participants behaviour, movement and that response will likely be natural and not be influenced by the researcher during the trail. Thus to understand the performance of the positioning system a novel real-world occupancy trial is designed in the form of a Lego brick model building task to ensure they are ecologically valid and imitates real world occupancy scenario in office spaces. Lego resources and model pictures were spatially distributed across the NGB floors in different rooms and corridors and non-expert participants with no previous idea of using the positioning system were recruited with their consent. The main objective of the Lego trial was to observe the robustness and accuracy of the positioning system within the metrics of accurate identification of occupancy of rooms and corridors during the transition from one zone to another. There were in total 22 participants from the NGB agreeing to participate in a total of 50 trials covering floor A and floor B.

3.1.2 Task Description

Lego building is used to simulate daily activity during occupancy and participants were asked to move between rooms and corridors within the NGB to collect materials and resources for the toy brick (Lego) building task. Each of the participants was provided with a smartphone with the data logger installed and running, copies of floor maps and instruction sheet. They had to complete as many Lego models as they can within 20 minutes from an existing list of 8 Lego model with the phone in their pocket. During the trial the researcher shadowed them while manually recording their time and location as well, creating a journey log in the process for validation. The recorded data from the smartphone is postprocessed to compute the participant's position showing their movement trajectory and the journey log is used to compare and validate their occupancy.

The participants had complete control over route choices to reach their destination and were not advised or distracted by the researcher during the task. The instruction was changed for different participants by changing the location of the resources to be collected and their starting point so that we can observe occupancy across a variety of locations and routes within the building. A sample instruction list is provided below.

- 1. You will start from a desk in room B05, floor-B (to be shown by the researcher). Keep the phone in your trouser pocket and wait for approximately 20 seconds for initialisation after the data logger is started by the researcher.
- 2. Go to Corridor 4 and Corridor 2 (marked on the floor map) and find the Lego model pictures on the floor. There will be a total of eight pictures.
- 3. Keep the Lego pictures with you as well as all other sheets.
- 4. Take the Lego pictures in the room A20 floor- A. The Lego bricks are also kept in room A20, marked in the map. You can check the Lego pictures decide which model you want to build and gather the Lego bag. You can then go and build Lego models in room A19, Floor A. You can only take one Lego picture for building at a time. If you finish or decide to change your mind and build another one, you can check the remaining Lego model pictures in room A20 and collect appropriate Lego structures as many times as you want, but you can build the models only in room A19.
- 5. You will have to complete as many models as you can within 20 minutes.
- 6. When finished, the researcher would ask you to stop, and he will stop the data logger app for you.

4. Results

The trial data are post-processed and few segments of occupancy with a lot of movements are illustrated to show the time series of computed occupancy against actual occupancy, recorded by the researcher when shadowing the participants. The plots in Figure 5 below helps to observe the accuracy of the monitored occupancy and shows the time lag present between the processed and the recorded occupancy. It can be seen that there are instances of misidentification of room occupancy during the trial as well as some time lag. To further investigate the time lag we observe few individual transitions from different trial segments and compare against the computed and recorded time. The time difference between the computed and recorded transitions can be observed in Figure 6 below. The maximum mean time difference is 3.55 second and a minimum of 1.142 seconds while the median shows a maximum of 3 seconds. The time lag between the computed occupancy and recorded occupancy show the latency the particle filter engine suffers when computing the position estimation. Many factors may come into effect when trying to identify the underlying cause such as missing step detection, step length overestimation and the quality of Wi-Fi and BLE radio signals and any delay in their reception and recording by the data logger due to hardware issue. We can see from the plots that the order of time lag is in seconds and not significantly higher.



Figure 5. Time series of computed occupancy against recorded occupancy from selected segments of trials.



(c)

(a) Difference Mean: 2.44 sec, Median: 3 sec



(d)

(b) Difference Mean: 1.14 sec, Median: 3 sec



(c) Difference Mean: 1.42 sec, Median: 2 sec

(d) Difference Mean: 3.55 sec, Median: 2 sec

Figure 6. Transition time difference when moving between rooms and corridors during occupancy trials.

From a navigation and positioning systems point of view, this is a drawback which may impact other external functionalities synchronised for the real-time update. Notably, application areas requiring precise time-specific occupancy information such as in smart buildings to identify occupied and unoccupied periods of rooms, corridors and stairways to adjust HVAC systems of the building for optimal energy usage and improved building energy efficiency. In such applications, the occupancy information will be ideally used to develop stochastic time series models using algorithms such as HMM in [8], [36]–[38] which should be able to tolerate up to a certain degree of time lag as long as it does not differ in the order of minutes and hours. The time lag can thus be adjusted and smooth out with some confidence level during the data pre-processing stage and create a probability distribution of the likelihood of building occupancy throughout the day of individuals and also develop aggregate observation for similar groups of occupants. As such we find the level of time lag in this positioning system acceptable for some selected applications of smart buildings related to occupancy detection and suitable for occupancy model development using the computed occupancy data.

5. Discussion

5.1. Performance Analysis

To evaluate the performance of the positioning system when it comes to detecting occupancy, it is crucial to identify some key performance indicator which can be used as a criterion to determine the performance metric. This criterion may include but not limited to the correct identification of rooms, corridors, and stairs the participant occupies in the course of their occupancy in the building. In some cases, the accuracy level can be compromised for zones comprising of multiple adjacent rooms and corridors when modelling building occupancy. Room level detection accuracy will help to identify the exact number of occupants similar to [22]–[24] if correctly identified. One of the most extensive validations of occupancy detection model is seen in the work of Page et al. in [16] with single-occupant office rooms and in [8], [39] with multiple occupants and multiple zones. In both the latter cases, the

researchers were developing occupancy models with zone level accuracy for state transition. As such we can set two levels of performance necessary for developing any BOM; room level detection accuracy and zone level detection accuracy

From our positioning solution, the continuous trajectory of occupancy has the potential to provide an apparent picture of an individual's occupancy and can also be fused with the standalone Wi-Fi/BLE observation to understand zone level occupancy like Figure 7. This is particularly true in situations of incorrect or partially correct occupancy sequence where we may not get a complete picture of the journey within room level accuracy but still manage to extract enough information to infer the overall occupancy pattern of the individual. These data can then be used to develop the occupancy model both at an aggregate and individual level, which is nearly impossible with the existing methodologies of occupancy detection techniques discussed earlier. Based on these understanding we can list a few selected performance indicator for room level accuracy used to measure the quality of the trials below.

- Occurrences of completely incorrect occupancy sequence caused when all or a significant portion of the route is computed incorrectly and not suitable for room level accuracy.
- Partially incorrect occupancy sequence shows a wrong occupancy route computed for a part of the task, but the particle filter output may realign with the correct route afterwards.
- Failure at initialization implies a failure to collect data from one or more sensors due to a hardware issue in the smartphone.
- Correct occupancy sequence implies position computed to reflect actual occupancy from the task and providing correct identification of rooms and corridors during the transition.



Floor A

Floor B

Figure 7. The floor plan of NGB and its partition into zones and rooms to understand occupancy.

These indicators are used to determine the outcome of the trials and evaluate the robustness of this positioning system when tested by non-expert users with different physical characteristics and motion traits; such as walking speed, height, body postures, and route choice. Even with failures to detect occupancy trajectory accurately all the time, zone level detection accuracy was appreciable which is what most of the other occupancy detection methods, discussed above tries to achieve. The occupancy trajectory even if partially incorrect can be combined with zone level occupancy trajectory and used to

interpret the most likely sequence of transitions through rooms and corridors applying appropriate machine learning algorithms and data analysis techniques. Thus demonstrating the overall potential to produce occupancy information at a greater detail compared to any existing occupancy monitoring technologies and techniques discussed earlier. The overall trial performance can be seen in Table 1 below for room level accuracy and Table 2 for zone level accuracy. Out of 50 trials, there were 5 completely failed trials due to a crashed data logger at initialisation, wrong heading estimate at initialisation due to mishandling of the phone in the pocket and a missed stairwell detection causing subsequent misidentification of floor map as seen in Table 1 below. In total, we found 25 trials with a correct estimation of every occupancy transition. If we consider combining partially incorrect occupancy estimations with zone level identifications to identify likely transitions, the total number of trials with correct occupancy estimation rises to 34 out of 45 successful trials. The room level detection including every transition through stairways and corridors shows an overall accuracy of 82.5% with a precision of 90% and a recall of 90%. The zone level detection shows an accuracy of 97.5 % with a precision of 94.9 % and a recall of 95.2%. The designation of the zones in this research was simply based on the major coverage areas of the occupancy trial and ease of demonstration purpose and can be changed as required.

Table 1. Lego	trial	room	level	accuracy	chart

Performance Indicator	Number of Trials			
Correct estimation of occupancy	25			
Failure at initialization /data logger crashed	1			
Incorrect occupancy estimation	9			
Partially incorrect occupancy estimation	9			
Missing stairway detection	1			
Failure due to wrong heading estimate	3			
Missing Wi-Fi/BLE signal	1			
Particle filter restart	1			
Total	50			

Table 2. Lego trial zone level detection chart corresponding to Figure 7

Trial Number	Zone Detection	Trial Number	Zone Detection
Trial 1	(Z3) Floor B, St3, (Z4, Z3, Z1*) Floor A	Trial 24	(Z3,Z2^)Floor B,St3,(Z4,Z3,Z1,Z3,Z1,Z3,Z1)Floor A
Trial 2	(Z1, Z2, Z3) Floor B, St3, (Z4, Z3, Z1*) Floor A	Trial 25	(Z3) Floor B, St3, (Z4, Z3, Z1, Z3, Z1) Floor A
Trial 3	(Z2) Floor B, St1, (Z1, Z2, Z1*) Floor A	Trial 26	(Z3) Floor B, St3, (Z4, Z3,Z1,Z3,Z1,Z3*^)Floor A
Trial 4	(Z1, Z2, Z3) Floor B, St3, (Z4, Z3, Z1*) Floor A	Trial 27	(Z2, Z3) Floor B, St3, (Z4, Z3, Z1, Z3, Z1) Floor A
Trial 5	(Z3) Floor B, St3, (Z4, Z2, Z1*) Floor A	Trial 28	(Z2^) Floor B, St2, (Z1, Z3, Z1, Z3, Z1) Floor A
Trial6	(Z1, Z2) Floor B, (Z1, Z4, Z3, Z1*) Floor A	Trial 29	(Z1, Z2) Floor B, St1, (Z1, Z3, Z1, Z3, Z1) Floor A
Trial 7	(Z1) Floor A, St2, Z1*	Trial 30	(Z2, Z3) Floor B, St3, (Z4, Z3, Z1, Z3, Z1) Floor A
Trial 8	(Z2, Z1) Floor B, St2, (Z1*) Floor A	Trial 31	(Z2) Floor B, St1, (Z1, Z3, Z1, Z3, Z1) Floor A
Trial 9	(Z2, Z1) Floor B, St2, (Z1*) Floor A	Trial 32	(Z2) Floor B, St1, (Z1, Z3, Z1, Z3, Z1) Floor A
Trial 10	(Z3, Z2) Floor B, St1, (Z1*) Floor A	Trial 33	(Z1, Z3 [^]) Floor B, St1, (Z1, Z3, Z1, Z3, Z1) Floor A
Trial 11	(Z3) Floor B, St1, (Z1*) Floor A	Trial 34	(Z2, Z1^) Floor B, St1, (Z1, Z3, Z1, Z3 Z1) Floor A
Trial 12	(Z1) Floor B, St2, (Z1*) Floor A	Trial 35	(Z3, Z2) Floor B, St1, (Z1, Z3, Z1*) Floor A
Trial 13	(Z3) Floor B, St3, (Z4, Z3, Z1*) Floor A	Trial 36	(Z3, Z2) Floor B, St1, (Z1, Z3, Z1, Z3, Z1) Floor A
Trial 14	(Z1, Z2) Floor A, St2, (Z1, Z3, Z1*) Floor A	Trial 37	(Z3, Z2) Floor B, St1, (Z1, Z3, Z1, Z1*^) Floor A
Trial 15	(Z1) Floor B, St2, (Z1, Z3, Z4, Z2, Z1*) Floor A	Trial 38	(Z2) Floor B, St1, (Z1,Z3,Z1,Z3,Z1,Z2,Z1) Floor A
Trial 16	(Z3) Floor B, St3, (Z4, Z3, Z4, Z2, Z1*) Floor A	Trial 39	(Z2, Z1^) Floor B, St1, (Z1,Z3,Z1,Z3,Z1) Floor A

Trial 17	(Z1) Floor B, St2, (Z1, Z3, Z4, Z2, Z1*) Floor A	Trial 40	(Z1, Z2, Z4, Z3, Z1, Z2, Z1, Z2, Z1) Floor B
Trial 18	(Z2, Z3) Floor B, St3, (Z4, Z3, Z1*) Floor A	Trial 41	(Z1, Z2, Z1, Z3, Z3^, Z4^, Z2^, Z1, Z2, Z1) Floor B
Trial 19	(Z2) Floor B, St1, (Z4, Z3, Z1*) Floor A	Trial 42	(Z1, Z2, Z1, Z3, Z1, Z2, Z1) Floor B
Trial 20	(Z3) Floor B, St3, (Z4, Z3, Z1*) Floor A	Trial 43	(Z1, Z2, Z1, Z3, Z1, Z2, Z1) Floor B
Trial 21	(Z2) Floor B, St3, (Z4, Z3, Z1*) Floor A	Trial 44	(Z1, Z2, Z1, Z3, Z1, Z2, Z1) Floor B
Trial 22	(Z3) Floor B, St3, (Z4, Z3,Z1*) Floor A	Trial 45	(Z1, Z2, Z1, Z3, Z1, Z2, Z1) Floor B
Trial 23	(Z3)Floor B, St3, (Z4, Z3, Z1*) Floor A		

Z = zone, St = stairway, * = movement within the same zone between multiple rooms, ^ = wrongly detected or missed zone

The chart in Table 3 below provides a comparison of our proposed IP solution, marked with an asterisk against a few widely used existing occupancy monitoring technologies, their advantages and limitations. We can see that the PIR sensor, environmental sensor and smart meter all fail to provide continuous tracking with various additional constraints such as cannot detect stationary state, time lag, dependence on environmental conditions and detection based on appliance usage only. The use of camera technology provides continuous tracking and ticks all the occupancy resolution features but is expensive to install and deploy and dependent on good illumination condition as well as high computational complexity of image processing and analysis. Thus we are left with Wi-Fi and BLE based solution providing zone level occupancy detection with proximity sensing but falling short of continuous tracking. We have extended it in our IP solution and incorporated PDR capability using smartphone inertial sensors to enable continuous active tracking of individuals. Thus providing identification of every possible sequence of occupancy transition throughout the building. Another important advantage of our system is since it can detect at an individual level we can generate personal movement profiles to understand occupancy patterns across office demographics as well. Although there are limitations related to radio map setup and possible latency in data transmission and reception from smartphone sensors.

Sensor		Occupanc	ey informat	ion resolutio	Cost	Privacy Issue	Advantage	Limitations	
	State	Quantity	Activity	Identity	Track	_	20540		
PIR	Yes	Yes	No	No	No	Low	No	Low	Cannot
								computation	detect static
								al cost	state
Environmental	Yes	Yes	Yes	No	No	Medium	No	Widely	Time lag and
(CO ₂ ,								available	sensitive to
Temperature)								and non-	environment
								intrusive	
Smart meter	Yes	No	Yes	No	No	No	Partial	Existing	Miss
								infrastructur	occupants
								е	not using
									appliance
Wi-Fi/BLE	Yes	Yes	No	No	Partial	Medium	Partial	Widely	Dependent
								available	on the

Table 3. Occupancy detection technology comparison chart, updated from source [17].

(Proximity	(Zone								smartphone
only)	level)								being turne
									on
Wi-Fi/BLE*	Yes	Yes	No	No	Continuous	Medium	Partial	Widely	Radio map
(Fingerprinting	(Room							available.	setup.
& multi-sensor	level							Detailed	Latency
integrated PDR	and							occupancy	logging dat
solution)	zone							eg:	from
	level)							transition	multiple
								between	sensors.
								rooms and	Missing dat
								corridors	due to sense
									blackout.
Camera	Yes	Yes	Yes	Yes	Yes	High	Yes	High	High
								Accuracy	computatio
									al
									complexity
									and
									illuminatio
									conditions

* = proposed IP solution in this paper

5.2. Limitations

The "extensive" multi-user trial showed promising results with the developed IP system based on their performance analysis. The positioning system proved capable of detecting occupancy of individuals both at room level and at zone level accuracy. Thus providing detailed transition information and route choices. The trials if extended over the entire day can provide daily occupancy patterns of the building. The errors in the positioning data for partially correct occupancy sequences can also be corrected to some extent by doing a post-process data analysis or improving the map matching technique. Map matching is an environmental feature matching technique widely used in positioning algorithms to aid the movement of objects or particles [40] to get a well-controlled projection of movements within the area at each epoch. The technique is already used in the implementation of the positioning system by incorporating floor plans; walls and doors as seen in Figure 1 and 2 and could be improved further by incorporating carefully selected landmarks to remove outliers and realign any unusual route with the more likely one. Machine learning algorithms such as Random Forests, Naïve Bayes and Neural Network can also be trained to develop models that can help predict more likely routes based on already detected transitions as input. It can help to identify and realign incorrectly detected or missing occupancy transitions. Although this would require the collection of additional data for training, testing and validation of the models. When it comes to the smartphone as a technology capable of providing positioning solution, the sensors are generally of inferior quality. Latest smartphones are likely to have sensors with better sensitivity and possibly improve the overall performance as such would be another area to explore further in the future and compare the accuracy between different models of smartphones. The sensors suffer from hardware latency, which may result in complete loss of data. Time synchronisation of all the sensor data is a big problem in the Android platform. Each sensor API has its time parameter when receiving and logging data. Another concern is the battery life of the android data logger. This kind of multisensory app running continuously throughout the day will be extremely power-hungry and will likely cause the battery to run out of charge within 8 to 10 hours. It would be ideal to develop a bespoke device with all the required sensors packed together like Intel Edison or Raspberry Pi but with a customisable scan rate of Wi-Fi, BLE, accelerometer, and gyroscope. There were some problems detecting floor transition with a few cases of missing stair detection. These lead to misidentification of floor map fingerprint database or significant gap in Wi-Fi or BLE data reception causing issues with data processing and in the process ending up with erroneous occupancy sequence or system crash. The overall accuracy could be further improved by increasing the number of BLE beacons and placing them carefully with the range modified to reduce overlapping areas of similar signal strength. The radio map needs to be updated from time to time to ensure any environmental changes such as changes in floor plan layout and furniture placement are taken into account. As such SLAM based radio mapping or crowdsourced update with the help of voluntary participants can be considered. The behaviour of participants also plays a significant role which can be unpredictable and challenging to model such as step detection, body posture, speed and angle of a turn at corridors. In short, the trials tried to ensure participants are at their most natural state when completing the task. Thus we can understand the array of navigation and positioning issues and related environmental factors that affect the accuracy of our positioning system. It will also help to identify what may be done in the future to mitigate or improve its accuracy.

5.3 Potential Applications

Typically we could get a detailed and accurate picture of the occupants and how they make use of the building from the occupancy data generated by the IP system. The data collected throughout the day could be analysed to understand not just occupancy patterns room-by-room or routes choices but also provide insight into the demographics of behavioural patterns of occupants at different times of the day and month specific to the type of building. These understanding can be applied for predictions in other similar buildings. It could also be used in the development of an improved occupancy or presence model such as the one stated in [16]. The model should have the potential to incorporate multiple states of occupancy as well as route choices and also include behavioural traits from different clusters of occupants when it comes to understanding building occupancy. Other applications could be in smart heating systems [41] that relies on the accurate prediction of building occupancy or as input to other behavioural and automation models [42]–[44] such a light switching, appliance usage, window opening

and blinds to name a few. Also, the potential to effectively manage energy consumption in buildings is very high by identifying the right policies and measures based on the occupancy data.

6. Conclusions

Research in the field of IP has made tremendous progress over the last decade and continuously being updated with innovations and improvements. This research looked to tap into that progress and demonstrated a novel contribution of IP and highlighted the limitations of some of the technology and techniques when used in this kind of real-world application. Unlike typical positioning trials in a controlled experimental setup, we have tried to recreate a real-world experience of building occupancy through our novel Lego-based task design and observe ecologically valid occupancy trial outcomes. The trials demonstrated the feasibility of carrying our occupancy monitoring through active tracking using our PDR based IP system. We found the level of accuracy appreciable and quite good when compared to existing occupancy detection systems with a lot of possibility for further improvements in future work. Compared to other methods for detecting building occupancy, IP provides information related to the sequence of occupancy at an individual level, which is unique and essential for developing multi-state presence model. It only needs to be set up once in any building during the PROBE stage and collect occupancy data from time to time to understand patterns of occupancy and demographics from willing participants. The Built Environment and Energy Efficient Buildings and Building Management community will benefit from the knowledge of another potentially effective method for detecting building occupancy. They will gain insight into how the positioning data has been used in this research and what other ways the data can be used in research areas related to understanding building performance. The data could help to develop more detailed occupancy models at an individual and aggregate level. The models can then be incorporated into building performance simulation to improve the prediction capability related to energy consumption and applied in smart home automation.

In general, from the research, we can conclude that a smartphone with its array of built-in sensors is a viable tool for the application of IP technology in our daily life. Wi-Fi and BLE are easy to deploy and cheap but may require additional aid when it comes to detecting unconstrained human movement in a natural environment. The inherent constraints embedded in a real-world scenario such as building occupancy is important to understand for any future improvement of the system related to the use of other IP technology and techniques. Nevertheless, it can be safely claimed that IP holds a lot of potential for the improvement of building occupant modelling. Future work suggestions would be to look into improvement of the particle filter algorithm or implementing a Kalman filter running in parallel to aid the positioning estimation. Identify the feasibility of using Wi-Fi in combination with BLE and or using more carefully placed BLE beacons alone. The motion detection algorithm could be made more responsive to non-walk like movement detection. A more stable step-length calculation technique could also be adopted. The accuracy of the positioning system could also be compared and evaluated against

Author Contributions: The lead author is responsible for the entire research which is part of his PhD. All the other authors provided significant guidance and advice in their capacity as academic supervisors.

Funding: The lead author, Md Shadab Mashuk is supported by the Horizon Centre for Doctoral Training at the University of Nottingham (RCUK Grant No. EP/L015463/1) and by the RCUK's Horizon Digital Economy Research Institute (RCUK Grant No. EP/G065802/1) and Satellite Applications Catapult (sa.catapult.org.uk/).

Acknowledgments: The lead author, Md Shadab Mashuk is grateful to all the Nottingham Geospatial Building staff and colleagues who supported in carrying out this research.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- C. F. Reinhart and C. Cerezo Davila, "Urban building energy modeling A review of a nascent field," *Building and Environment*, vol. 97. Pergamon, pp. 196–202, 15-Feb-2016, doi: 10.1016/j.buildenv.2015.12.001.
- [2] A. C. Menezes, A. Cripps, D. Bouchlaghem, and R. Buswell, "Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap," *Appl. Energy*, vol. 97, pp. 355–364, Sep. 2012, doi: 10.1016/J.APENERGY.2011.11.075.
- W. F. E. Preiser, "Post-occupancy evaluation: How to make buildings work better," *Facilities*, vol. 13, no. 11, pp. 19–28, Nov. 1995, doi: 10.1108/02632779510097787.
- [4] H. N. Rafsanjani, C. R. Ahn, and M. Alahmad, "A review of approaches for sensing, understanding, and improving occupancy-related energy-use behaviors in commercial buildings," *Energies*, vol. 8, no. 10.
 MDPI AG, pp. 10996–11029, 01-Oct-2015, doi: 10.3390/en81010996.
- J. Yang, M. Santamouris, and S. E. Lee, "Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings," *Energy Build.*, vol. 121, pp. 344–349, Jun. 2016, doi: 10.1016/j.enbuild.2015.12.019.
- [6] M. S. Mashuk, J. Pinchin, P.-O. Siebers, and T. Moore, "A smart phone based multi-floor indoor

positioning system for occupancy detection," in 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS), 2018, pp. 216–227, doi: 10.1109/PLANS.2018.8373384.

- J. Chapman, P. O. Siebers, and D. Robinson, "On the multi-agent stochastic simulation of occupants in buildings," *J. Build. Perform. Simul.*, vol. 11, no. 5, pp. 604–621, 2018, doi: 10.1080/19401493.2017.1417483.
- [8] Z. Chen and Y. C. Soh, "Modeling building occupancy using a novel inhomogeneous Markov chain approach," in *IEEE International Conference on Automation Science and Engineering*, 2014, vol. 2014-Janua, pp. 1079–1084, doi: 10.1109/CoASE.2014.6899459.
- [9] B. Ai, Z. Fan, and R. X. Gao, "Occupancy estimation for smart buildings by an auto-regressive hidden Markov model," in *Proceedings of the American Control Conference*, 2014, pp. 2234–2239, doi: 10.1109/ACC.2014.6859372.
- [10] L. M. Candanedo and V. Feldheim, "Accurate occupancy detection of an office room from light, temperature, humidity and CO2 measurements using statistical learning models," *Energy Build.*, vol. 112, pp. 28–39, 2016, doi: 10.1016/j.enbuild.2015.11.071.
- [11] Z. Yang, B. Becerik-Gerber, N. Li, and M. Orosz, "A systematic approach to occupancy modeling in ambient sensor-rich buildings," *Simulation*, vol. 90, no. 8, pp. 960–977, Aug. 2014, doi: 10.1177/0037549713489918.
- [12] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy Estimation Using Only WiFi Power Measurements," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1381–1393, 2015, doi: 10.1109/JSAC.2015.2430272.
- [13] Y. Wang and L. Shao, "Understanding occupancy pattern and improving building energy efficiency through Wi-Fi based indoor positioning," *Build. Environ.*, vol. 114, pp. 106–117, Mar. 2017, doi: 10.1016/j.buildenv.2016.12.015.
- [14] T. Zhang, P.-O. Siebers, and U. Aickelin, "Modelling Electricity Consumption in Office Buildings: An Agent Based Approach," May 2013.
- [15] D. Wang, C. C. Federspiel, and F. Rubinstein, "Modeling occupancy in single person offices," *Energy Build.*, vol. 37, no. 2, pp. 121–126, 2005, doi: 10.1016/j.enbuild.2004.06.015.
- [16] J. Page, D. Robinson, N. Morel, and J.-L. Scartezzini, "A generalised stochastic model for the simulation

of occupant presence," *Energy Build.*, vol. 40, no. 2, pp. 83–98, Jan. 2008, doi: 10.1016/j.enbuild.2007.01.018.

- [17] L. Rueda, K. Agbossou, A. Cardenas, N. Henao, and S. Kelouwani, "A comprehensive review of approaches to building occupancy detection," *Build. Environ.*, vol. 180, p. 106966, Jun. 2020, doi: 10.1016/j.buildenv.2020.106966.
- [18] M. S. Zuraimi, A. Pantazaras, K. A. Chaturvedi, J. J. Yang, K. W. Tham, and S. E. Lee, "Predicting occupancy counts using physical and statistical Co2-based modeling methodologies," *Build. Environ.*, vol. 123, pp. 517–528, 2017, doi: 10.1016/j.buildenv.2017.07.027.
- B. Dong *et al.*, "An information technology enabled sustainability test-bed (ITEST) for occupancy detection through an environmental sensing network," *Energy Build.*, vol. 42, no. 7, pp. 1038–1046, Jul. 2010, doi: 10.1016/j.enbuild.2010.01.016.
- [20] A. Corna, L. Fontanab, A. A. Nacci, and D. Sciuto, "Occupancy Detection via iBeacon on Android Devices for Smart Building Management," in *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2015, 2015, pp. 629–632, doi: 10.7873/DATE.2015.0753.
- [21] A. A. Adamopoulou, A. M. Tryferidis, and D. K. Tzovaras, "A context-aware method for building occupancy prediction," *Energy Build.*, vol. 110, pp. 229–244, 2016, doi: 10.1016/j.enbuild.2015.10.003.
- [22] R. H. Dodier, G. P. Henze, D. K. Tiller, and X. Guo, "Building occupancy detection through sensor belief networks," *Energy Build.*, 2006, doi: 10.1016/j.enbuild.2005.12.001.
- [23] T. Labeodan, W. Zeiler, G. Boxem, and Y. Zhao, "Occupancy measurement in commercial office buildings for demand-driven control applications - A survey and detection system evaluation," *Energy* and Buildings. 2015, doi: 10.1016/j.enbuild.2015.02.028.
- [24] M. Jia, R. S. Srinivasan, and A. A. Raheem, "From occupancy to occupant behavior: An analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency," *Renewable and Sustainable Energy Reviews*, vol. 68. pp. 525–540, 2017, doi: 10.1016/j.rser.2016.10.011.
- [25] F. Haldi and D. Robinson, "Interactions with window openings by office occupants," *Build. Environ.*, vol. 44, no. 12, pp. 2378–2395, 2009, doi: 10.1016/j.buildenv.2009.03.025.
- [26] F. Haldi and D. Robinson, "Adaptive actions on shading devices in response to local visual stimuli," J.

Build. Perform. Simul., vol. 3, no. 2, pp. 135–153, Jun. 2010, doi: 10.1080/19401490903580759.

- [27] C. F. Reinhart, "Lightswitch-2002: a model for manual and automated control of electric lighting and blinds," *Sol. Energy*, vol. 77, no. 1, pp. 15–28, 2004, doi: 10.1016/j.solener.2004.04.003.
- [28] X. Dai, J. Liu, and X. Zhang, "A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings," *Energy and Buildings*, vol. 223. Elsevier Ltd, 15-Sep-2020, doi: 10.1016/j.enbuild.2020.110159.
- [29] R. Mautz, "Indoor positioning technologies," 2012, doi: 10.3929/ETHZ-A-007313554.
- [30] R. F. Brena, J. P. García-Vázquez, C. E. Galván-Tejada, D. Muñoz-Rodriguez, C. Vargas-Rosales, and J.
 Fangmeyer, *Evolution of Indoor Positioning Technologies: A Survey*, vol. 2017. 2017, pp. 1–21.
- [31] A. Basiri, P. Peltola, P. Figueiredo e Silva, E. S. Lohan, T. Moore, and C. Hill, "Indoor positioning technology assessment using analytic hierarchy process for pedestrian navigation services," in 2015 International Conference on Location and GNSS (ICL-GNSS), 2015, pp. 1–6, doi: 10.1109/ICL-GNSS.2015.7217157.
- [32] A. Basiri, E. S. Lohan, Pedro Figueiredo e Silva, P. Peltola, C. Hill, and T. Moore, "Overview of positioning technologies from fitness-to-purpose point of view," in *International Conference on Localization and GNSS 2014 (ICL-GNSS 2014)*, 2014, pp. 1–7, doi: 10.1109/ICL-GNSS.2014.6934176.
- [33] C. Andrade, "Internal, external, and ecological validity in research design, conduct, and evaluation," *Indian J. Psychol. Med.*, vol. 40, no. 5, pp. 498–499, 2018, doi: 10.4103/IJPSYM_J34_18.
- [34] M. R. Banaji and R. G. Crowder, "Some Everyday Thoughts on Ecologically Valid Methods," Am.
 Psychol., vol. 46, no. 1, pp. 78–79, 1991, doi: 10.1037/0003-066X.46.1.78.
- [35] M. A. Schmuckler, "What is Ecological Validity? A Dimensional Analysis," *Infancy*, vol. 2, no. 4, pp. 419–436, 2001, doi: 10.1207/S15327078IN0204_02.
- [36] J. R. Dobbs and B. M. Hencey, "Model predictive HVAC control with online occupancy model," *Energy Build.*, vol. 82, pp. 675–684, 2014, doi: 10.1016/j.enbuild.2014.07.051.
- [37] C. Wang, D. Yan, and Y. Jiang, "A novel approach for building occupancy simulation," *Build. Simul.*, vol. 4, no. 2, pp. 149–167, Jun. 2011, doi: 10.1007/s12273-011-0044-5.
- [38] S. H. Ryu and H. J. Moon, "Development of an occupancy prediction model using indoor environmental data based on machine learning techniques," *Build. Environ.*, vol. 107, pp. 1–9, Oct. 2016, doi:

10.1016/j.buildenv.2016.06.039.

- [39] C. Liao, Y. Lin, and P. Barooah, "Agent-based and graphical modelling of building occupancy," J. Build.
 Perform. Simul., vol. 5, no. 1, pp. 5–25, Jan. 2012, doi: 10.1080/19401493.2010.531143.
- [40] P. D. (Paul D. Groves, *Principles of GNSS, inertial, and multisensor integrated navigation systems*. Artech House, 2013.
- [41] V. Becker, W. Kleiminger, V. C. Coroamă, and F. Mattern, "Estimating the savings potential of occupancy-based heating strategies," *Energy Informatics*, vol. 1, no. S1, pp. 11–12, 2018, doi: 10.1186/s42162-018-0022-6.
- [42] T. Hong, S. C. Taylor-Lange, S. D'Oca, D. Yan, and S. P. Corgnati, "Advances in research and applications of energy-related occupant behavior in buildings," *Energy and Buildings*, vol. 116. pp. 694– 702, 2016, doi: 10.1016/j.enbuild.2015.11.052.
- [43] C. M. Clevenger and J. Haymaker, "The Impact Of The Building Occupant On Energy Modeling Simulations," Jt. Int. Conf. Comput. Decis. Mak. Civ. Build. Eng., pp. 1–10, 2006, doi: 10.1061/(ASCE)CP.1943-5487.0000323.
- [44] F. Haldi and D. Robinson, "The impact of occupants' behaviour on building energy demand," J. Build.
 Perform. Simul., vol. 4, no. 4, pp. 323–338, Dec. 2011, doi: 10.1080/19401493.2011.558213.