A BLUETOOTH BASED OCCUPANCY DETECTION FOR BUILDINGS

June Young Park, Thomas Dougherty, and Zoltan Nagy
Intelligent Environments Laboratory
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin, Austin, TX, USA

ABSTRACT

As occupants and their behaviors are responsible for a significant share of the energy consumption in buildings, it is of importance to gather occupancy data. In fact, gathering occupancy data is considered as one of the grand challenges in building information modeling. Capitalizing on the pervasiveness of mobile devices with Bluetooth (BT) functionality, in this paper, we propose an occupant detection system that uses the BT signal to infer occupant presence. We present the low-cost hardware based on the Raspberry Pi and the open-source software. We apply our approach in a real building environment with two experimental scenarios: 1) Occupancy estimation of a whole building level, and 2) Characterization of occupant types in a shared office. We estimate the ratio $r$ of detected BT devices to actual number of people to be $r \approx 0.64$. Our results show robust detection of occupants, as well as successful characterization of occupancy types as stationary, regular occupants, and visitors. Our method can be deployed quickly, and does not require the occupants to install a specific software. Thus, the proposed approach is especially useful for retrofit solutions.

Introduction

In recent years, the role of occupants and their behavior has been studied extensively, e.g., within the International Energy Agency’s Energy in Buildings and Communities program (Annex66). It has been concluded that occupants and their behaviors are responsible for a significant share of the energy consumption in buildings. For example, Shen et al. reviewed optimal control systems of office buildings using occupancy related data. These occupancy based building control approaches save 20-50% compared to conventional building control strategies (Shen, Newsham, and Gunay 2017). However, occupant related information is typically not the main consideration for building control researchers (Park and Nagy 2018). Further-more, over 100 academic and industry practitioners re-port that acquiring occupant related information is one of the grand challenges in building information modeling industry (Leite et al. 2016). In other words, operating building systems most efficiently requires occupancy information. Gathering occupant presence, however, is challenging, and the main topic of this paper.

Related research on occupancy detection

One can distinguish between direct and indirect approaches to detect occupant presence: the direct systems recognize the human as an object directly, while the indirect methods infer occupancy by measuring environmental changes from human activity. An ideal occupancy detection method is not intrusive to the occupants, i.e., no special devices or actions are required by the person.

The most common method for direct occupancy detection is using passive infrared (PIR) sensors, i.e., motion detectors (Nagy et al. 2015; Kim, Moon, and Yoon 2017). This sensor detects motion through the changes of the infrared radiation on its surface. Even though PIR sensors are very common in modern buildings, they are far from perfect. The main drawback is that they are detecting motion, and not presence. In other words, if a person is not moving, e.g., working at a computer or reading, the sensor will send a false negative signal, indicating that the room is vacated. To counteract this, typically PIR sensors are set up such that the signal for an empty room is only sent after a certain time threshold of non-activity has been reached, e.g., after 10min (Nagy et al. 2015). While this helps to reduce the false negatives, it also increases the false positives, i.e., it declares a room as occupied, when in fact the person has left. As a result, energy may be wasted. Another disadvantage is that PIR sensor needs a direct line of sight, which means the performance is highly dependent on the room geometry and furniture locations (Shen, Newsham, and Gunay 2017).

In contrast, indirect occupancy detection methods sense the change in the environment caused by occupants. For example, a sedentary human generates about 0.3L/min of carbon dioxide (CO₂). Due to this natural phenomenon, CO₂ sensor has been proposed for occupancy detection in buildings (Pedersen, Nielsen, and Petersen 2017). Rather than only using CO₂ sensor itself, researchers have combined with other environmental sensors as well (e.g., carbon-monoxide (CO), total volatile organic compounds (TVOC), small particles (PM2.5), acoustics, illumina-
Occupancy detection using Bluetooth

In this work, we use discoverable BT devices to infer occupancy in various realistic office settings. In their comparative study of multiple wireless protocols, Lee et al. argued that the advantages of BT over other technologies are robustness, low power consumption, quick deployment, and low cost (Lee, Su, and Shen 2007). So far only a few examples have been implemented in the built environment. Zhao et al. used BT to collect ground truth of occupancy data (Zhao et al. 2014). Their approach required pre-assigned, e.g., paired, BT devices, which limits the applicability in a general setting. On the other hand, Conte et al. developed a learning based occupancy detection system by modifying the iBeacon protocol which is a BT based indoor proximity system (Conte et al. 2014). Even though they achieved 84% accuracy, the experiment was only conducted in a laboratory setting, and the learning algorithm required to install special smart phone applications and manual human input.

Contribution

From these previous studies, we identified two major challenges for occupancy detection: 1) Deployment requirements: deploying environmental sensors is sensitive to physical conditions of room (i.e., window, diffuser, human location), while BT based approaches often require the occupants to install special smart phone application or use on additional device. 2) Applicability: the occupancy data can be utilized in various ways. However, the implementation of this information in a building management system is still limited due to the lack of IT infrastructure, especially in existing buildings. The contribution of this paper is twofold. First, we present the hardware for a BT based occupancy detection system. It is an easily-deployable and low-cost method for

Figure 1: A: Raspberry Pi Zero W, B & C: System deployment

ing a Raspberry Pi microcomputer. It can be readily used by researchers and practitioners who want to acquire occupancy data. Second, we apply our method in a real building environment, and evaluate its performance: 1) the number of occupants are estimated, and 2) occupant types are characterized in shared offices.

To address the deployment challenge, our proposed method only needs a power outlet for operation (possibly replaced by a battery), and it is a non-intrusive approach, which does not require occupants to install additional applications on their devices. In terms of the applicability challenge, we provide an open source IoT platform to reproduce, store, transfer and utilize various building occupancy data for different applications.

Methodology

Occupancy detection based on Bluetooth

Our approach relies on the fact that in today’s world, many people have a mobile phone with potentially enabled BT functionality. Further, many other BT devices besides mobile phones are in use, e.g., wireless headphones, speakers, cars, watch, etc. Similarly to the WiFi approach in the previous section, we can infer the presence of a person by detecting the presence of a BT signal. Each BT device has a media access control (MAC) address, which is a unique identifier assigned for networking communication purposes. We implement scanning mechanisms in our occupancy detection system: Scan is aiming for all the devices within the searchable range. Although not all BT devices are discoverable, most devices advertise their availability for connection, if the BT setting is activated. We use the Raspberry Pi (RPI) Zero W ($10 as of 2018) to execute our occupancy detection system. It is a small sized computer (Fig. 1) that has built-in BT capability. The software for the occupancy detection is built upon the Linux command, lscans (Blum 2008). The software is available for download and extension in our online repository, and provides an easy-to-use environment for building researchers who need occupancy data.
In addition, we activate the Wifi networking capabilities on each device to allow for scalable deployment: users i.e., researchers & practitioners, can access data remotely, and the software on the RPis can be automatically updated via a web-based repository. As a result, large amounts of diurnal occupancy related data can be handled efficiently.

Experimental scenarios
To investigate the proposed occupancy detection system, we develop experimental scenarios based on occupant related information. The test-bed for the experiments is the Ernest Cockrell Jr. Hall (ECJ) on the main campus of The University of Texas at Austin (Fig. 2). The ECJ building has a gross area of roughly 240,000ft$^2$ (=22,300m$^2$) and contains various usage types (e.g., faculty offices - individual offices, graduate students offices - shared offices, laboratories, and mid and large size classrooms).

As summarized in Fig. 3, the experiments vary by the number of people, that we wish to detect, resulting in different space types and potential applications. We vary these occupancy types from a single person to a crowd of people (< 150 persons): 1) With a large group of people, estimating occupancy level in a building, and 2) In detail, characterizing the occupant types in a shared office.

Validation for inference ratio
We deploy our system in mid and large sized classrooms (Fig. 4) for five weeks (09/09/2017 - 10/13/2017), where we have a recurring occupancy pattern due to academic schedules (see Tab. 1). On that schedules, we manually count the number of people at the beginning of each lecture. Having the ground truth information, the ratio between the number of advertising BT devices and the actual number of people is statistically inferred.

In this defined setting, the ratio ($r$) of discoverable BT devices ($N_{BT}$) to the actual number of people ($N_p$) was calculated. While $N_{BT}$ was acquired by the detection system, we manually counted the ground truth $N_p$ in the classrooms. The confidence interval for $r$ is calculated as,

$$r = \frac{N_{BT}}{N_p} \in r - t_{n-1} \times \frac{S}{\sqrt{n}} + r + t_{n-1} \times \frac{S}{\sqrt{n}}$$

where $r$ and $S$ are sample mean and standard deviation, respectively, $n$ is the sample size of our data collection, and $t_{n-1} = 2.57$ for the 99% confidence interval for sufficiently large $n$. The direct application of this experiment is the validation of of BT based occupancy detection. In addition, it allows to estimate $r$ for a university student population, which can then be used for inferring the actual number of people by detected BT devices.

Building occupancy estimation
We then installed the proposed system to count the number of BT enabled devices at the entrances of the
Table 1: Classroom schedules for validation experiment

<table>
<thead>
<tr>
<th>Classroom</th>
<th>Day</th>
<th>Start Time</th>
<th>End Time</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Tue, Thu</td>
<td>14:00</td>
<td>15:15</td>
<td>45-56</td>
</tr>
<tr>
<td>C2</td>
<td>Tue, Thu</td>
<td>12:30</td>
<td>14:00</td>
<td>32-39</td>
</tr>
<tr>
<td>C3</td>
<td>Tue, Thu</td>
<td>14:00</td>
<td>15:15</td>
<td>24-30</td>
</tr>
<tr>
<td>C4</td>
<td>Mon, Wed, Fri</td>
<td>14:00</td>
<td>15:50</td>
<td>13-21</td>
</tr>
</tbody>
</table>

Table 2: Entrance points in the ECJ building

<table>
<thead>
<tr>
<th>Entrance</th>
<th>Location</th>
<th>Orientation</th>
<th>Connection to/from</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>Main lobby</td>
<td>West</td>
<td>pedestrian entry</td>
</tr>
<tr>
<td>E2</td>
<td>Elevator</td>
<td>East</td>
<td>garage</td>
</tr>
<tr>
<td>E3</td>
<td>Secondary lobby</td>
<td>West</td>
<td>pedestrian entry</td>
</tr>
<tr>
<td>E4</td>
<td>Staircase</td>
<td>North</td>
<td>bus stop</td>
</tr>
</tbody>
</table>

ECJ building. The data was collected for five weeks (09/09/2017 - 10/13/2017) with one minute resolution. There are four entrances to the ECJ building (Tab. 2, floor-plans not shown). By measuring all these points in the same time periods, we can also compare the crowdedness of different entrances of the building. For the next experiment, we decreased the scale to the shared offices. More detail occupant related information is discovered.

Occupant type characterization

Using the lescan command, the system counts how many advertising BT devices are present in a shared office with one minute resolution, and stores the timestamp and their MAC address. In this experiment, we firstly attempt to infer the typical occupancy rate in four shared offices (Fig. 5), O1 - O4 based on the number of detected BT devices during five weeks (09/09/2017 - 10/13/2017). Each office is mainly occupied by six graduate students, and other people can visit the offices for multiple activities (i.e., short & long term meetings). The number of BT devices is normalized by the maximum number of BT devices in each office, and averaged hourly to compare to the ASHRAE 90.1 Appendix G.

In addition, since the MAC addresses of the occupants are not known, we wish to characterize the occupancy data automatically based on frequency of occurrence of the devices. The probability of occurrence is calculated based on the likelihood of presence during the experimental period. For each BT device, the system logged the number of occurrences in every minute. We then normalized this by the total minutes of the experiment (35 days x 24 hours/day x 60 min/hour). Then, we apply three ranges to classify the presence probability of each device. We expect to have high probability ($P \geq 0.375$) for stationary devices that are always in the office, medium probability ($0.1 \leq P < 0.375$) for regular occupants and low probability ($0.01 \leq P < 0.1$) for visitors. Very low probabilities of occurrence ($P < 0.01$) for unique events are neglected. The threshold values for the probabilities are determined based on usability pattern in these offices: 0.1, 0.01 and 0.375 correspond approximately to 1.5, 17 and 63 hours in one typical workweek, respectively. For other buildings, these probability ranges can be adapted.

Results

The inference ratio

Each box plot represents the distributions of the number of discovered BT devices by lectures in the classrooms, and the red dots highlights the actual number of people (Figure 6). The numbers of discovered BT devices vary due to the fluctuation of people at the beginning and end of lectures. Compared to Tab. 1, there are fewer devices than the actual number of people in C1 & C2, while the system detects more devices than the number of occupants in the classrooms C3 & C4. This is mainly because, the classrooms C3 & C4 are smaller than the others, and surrounded by other rooms and corridors (Fig. 4). As a consequence, it is likely that they captured BT devices outside of the classrooms (i.e., other classrooms, corridor). How-ever, the locations of the classrooms C1 & C2 are rather isolated, and the room sizes are large enough that the system captured the devices which are only in the classrooms. To correct for this flaw in the experimental design, we only used the data from the classrooms C1 & C2 in our further analysis. From $n = 1143$ samples, we determine
Figure 6: Distribution of the number of detected BT devices by classrooms (the actual numbers of people are presented by the red dots)

The mean of the ratio $\overline{r} = 0.65$, and the standard deviation $s = 0.17$. Finally, we find the 99% confidence interval for $r$ as

$$r = \frac{N_{BT}}{N_p} \in [0.637, 0.663] \quad (2)$$

This indicates that if there are sufficiently many people in an academic building, 100 people will possibly have at least 64, and at most 66 discoverable BT devices (with 99% confidence).

We now move to apply the ratio $r$ to calculate the actual number of people at the entrances in the ECJ building.

Whole building occupancy estimation - entrances

Fig. 7 illustrates the number of BT devices around all the entrances in the building, each row represents daily profiles at different entrances. The daily profiles were calculated by averaging the number of BT devices throughout the experiment period. We find that E1 (main lobby) is the busiest area, while E2 is the least busy area, likely because this entrance is solely used as connection to the garage and very few people commute by car. The other entrances (E3 & E4) show similar peak periods and profiles on each day. This is because both entrances are close to most of the classrooms in the building and courses have Monday & Wednesday and Tuesday & Thursday repetitive schedules. In general, there are increasing patterns around 10 am for weekdays, which indicates that the building is increasingly occupied before noon, and the number of devices is decreasing gradually with a longer timespan compared to the increasing patterns.

Applying the conversion factor (0.64 - 0.66) calculated in previous section, we can infer the magnitude of occupancy level during peak times. Tab. 3 describes the peak occupancy levels for each entry point. On Wednesday around 10 am, a total of 249 BT devices are detected at the entrances E1, E3, and E4, which is equivalent to the minimum of 378 and the maximum of 388 people in the entrances. Since it is morning time, we assume that most people are coming into the building for classes and work. While the actual number of people can only be estimated, we hypothesize that this approach is sufficient to understand temporal features of the building, when the sensors are placed in the vicinity of the entrances of the building. In particular, dynamic features, i.e., peak occupancy times and weekday patterns of buildings are crucial information for establishing smart energy management among multi-ple buildings. In addition, we can easily implement the
system outputs for demand driven building control schedules (Nagy, Vazquez-Canteli, and Park 2018).

Occupant type characterization - shared offices

We infer the occupancy rate via the number of discovered BT devices. In Fig. 8, the four top rows show the daily occupancy rates for the offices (O1 - O4), while the last row shows the ASHRAE standard used as energy simulations input for office type (ASHRAE 90.1 Appendix G). Apparently, the profiles are different from the ASHRAE standard, and that each office has its own distinct daily profiles. There are three implications of this observation: (1) As building energy modelers refer to the ASHRAE standards for the occupancy rate, having true occupancy data can be used as simulation input for existing buildings to dramatically improve simulation accuracy. (2) Different offices show different occupancy rates, therefore a central-ized control strategy (fixed set-points) is not necessarily appropriate to satisfy the loads in different rooms. (3) The occupancy schedule is not same even in the same office by day. For example in the shared office O4, the peak occupancy level is only captured around noon on Tuesdays. This indicates that the building control strategy should be adapted for dynamic daily occupancy patterns to accomplish both energy saving and occupant comfort. 

Next, Figure 9 summarizes the occupant types for the shared offices. Each office had a few (< 5) stationary devices, e.g., desktop computer, printer, etc. Next, we found 5, 0, 4, and 2 regular occupants in each office, respectively. They regularly occupied the offices at least 17 hours ($P > 0.1$) and at most 63 hours ($P < 0.375$) in a typical workweek. In addition, we found around 10 - 30 visitors who occupied these shared offices between 1.5 hours ($P > 0.01$) and 17 hours ($P < 0.1$) in a typical workweek. Since not all the occupants activate their BT devices and one person might have multiple active BT devices, it is challenging to generalize that the numbers in Fig. 9 correspond to the actual number of occupants. However, we can further use this information for developing an occupant centered control strategy. For example, a controller should aim to provide comfort for the regular occupants rather than for visitors. When multiple people share same office space, the building control system needs to balance the preferences of multiple occupants (Wilson 2015), and characterizing occupant types in a shared office is essential to prioritizing control strategy for proper occupants. This application stands for a paradigm shift for the development of occupant centered, i.e., personalized control systems (Nagy et al. 2015; Nagy, Yong, and Schlueter 2016). In addition, as our proposed method is an efficient way to acquire true occupancy data, it can be further used as simulation input to calibrate building simulation models, especially when analyzing retrofit scenarios for existing buildings.

Discussion

As introduced in the beginning of the paper, we identified two major challenges for current occupant detection systems. First, current systems require special installation conditions. This drawback is a further barrier to retrofit existing buildings. In the U.S., the retrofitting rate of existing buildings is approximately 2.2% per year, and most of retrofitting strategies are return-on-investment analysis for replacing lighting or HVAC equipment (Olgyay and Seruto 2010). However, integrated measures for building retrofit including upgrading building control system and analyzing occupant behavior contributes energy savings up to 50% (Killien 2011). In particular, our proposed system is a suitable approach for measuring occupant related information in existing buildings. This is because, our system consists of open source platform, which is low cost and easily deployed: an electrical power outlet is the only requirement, and there is no additional installation on the occupant side.

Another challenge is the lack of applicability for building control system. Based on our reviews, most of the research on this topic covers a single application. Our approach provides an extensible platform for various future applications in building control system by, e.g., with Wifi networking, the acquired occupant related data can be integrated into the existing building automation systems (e.g., KNX, BACnet, Zigbee) (Jung et al. 2013). Recently, the Bluetooth developer consortium announced
that the upcoming version of BT will support mesh networking. A system with this new BT mesh networking feature provides a structure which potentially can count all the searchable BT devices in a building. In other words, all BT devices can provide occupant related information to existing building automation systems with minimal WiFi network.

One limitation of our approach is that the system relies on the occupants to have a BT enabled personal device that they carry with them. If an occupant does not activate the BT function or bring a mobile phone, the system will miss one occupant. Still, the potential is significant: In 2007, 6 BT devices were found for every 100 people in an urban environment (Nicolai and Kenn 2007). In this work, we identified ≈6.5 BT devices for every 10 people in a university building. Given that more than 3.4 billion products are manufactured with BT technology every year, and there are various BT devices, e.g., mobile phone, laptop computer, headphone, smart bracelet, and chain, it is safe to assume that in the near future BT will be a promising technology for occupancy detection in buildings.

In the future, re-calibration of the ratio \( r \) between the number of BT enabled devices and actual people may be necessary. Specifically, \( r \) may vary in buildings with other use types due a different occupant type (researchers vs students, etc) who may have a different number of BT enabled devices. In addition to recalibration for various building types, we will re-calibrate \( r \) in the ECJ building over time, i.e., every 6 months, to reflect the potentially increased number BT devices in our daily life. In practice, one has to determine a practical and economical recalibration procedure.

A second limitation is that our system only counts the number of BT devices, the direction of motion of the BT device/occupant, is not monitored. For example, in Fig. 7, we assumed that most people are coming and leaving during the morning, and afternoon, respectively. A university building generally follows this assumption because of the academic schedule, but it is difficult to generalize to other buildings. A denser sensor network could properly track the flow of occupants. For instance, if one sensor node logs a device and that same device is captured later at another sensor node, the system can record the temporal flow of this device in the building. This could then be expanded to urban scale to understand pedestrian motion patterns for urban planning purposes.

Finally, privacy concern is an important issue in IoT applications: since the proposed system is passive from occupants perspective, it is necessary to protect occupants’ information. The MAC address itself is not directly related to personal information, and one can further de-identify it using, e.g., the MD5 hashing technique (Deepakumara, Heys, and Venkatesan 2001). Updating the seed for the hash regularly, e.g., hourly, further minimizes the possibility to track a specific device throughout a building.

**Conclusion**

In this paper, we evaluated the potential of BT technology for occupancy detection. We described two experimental scenarios and evaluated it in a real building environment: 1) Validating with ground truth data, we found the ratio to estimate building occupancy at the building entrance points. 2) The system also characterizes occupant type by a likelihood of presence. To conclude, understanding and actively utilizing occupant related information in a building is imperative. In implementing our proposed open source detection system, researchers and practitioners can acquire occupant related information, which can be used, e.g., for occupant centered control, or simulation input.

---

**Figure 9: Number of BT devices by presence probability:** stationary device (\( P > 0.375 \)), regular occupant (\( 0.1 < P < 0.375 \)), and visitor (\( 0.01 < P < 0.1 \)) (limits are represented by the vertical dashed lines; Average work hours per week in the U.S. (38.6hrs) are described by the vertical red lines).
REFERENCES


Conte, Giorgio, Massimo De Marchi, Alessandro Antonio Nacci, Vincenzo Rana, and Donatella Sciuto. 2014. “BlueSentinel: a first approach using iBeacon for an energy efficient occupancy detection system.” *BuildSys@SenSys*. 11–19.


