

WHERE TO GO? SMART GUIDANCE BASED ON IOT SENSOR-DATA

Robin Effing¹, Robert J. Brouwer², Asen Iliev², Thomas Teunissen² and Fons Wijnhoven²

¹*Saxion University of Applied Sciences / University of Twente, P.O. Box 70.000, 7500 KB Enschede, The Netherlands*

²*University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*

ABSTRACT

There is an increasing interest in indoor occupation and guidance information for business and societal purposes. Scientific literature has paid attention to various ways of detecting occupation using different sensors as data source including various algorithms for estimating occupation rates from this data. Gaining meaningful insights from the data still faces challenges because the potential benefits are not well understood. This study presents a proof-of-concept of an indoor occupation information system, following the design science methodology. We review various types of sensor data that are typically available or easy-to-install in buildings such as offices, classrooms and meeting rooms. This study contributes to current research by incorporating business requirements taken from expert interviews and tackling one of the main barriers for business by designing an affordable system on a common existing infrastructure. We believe that occupation information systems call for further research, in particular also in the context of social distancing because of covid19.

KEYWORDS

Occupation Detection, Occupation Capacity, Internet of Things, Smart Surveillance, WiFi Tracking, Smart City

1. INTRODUCTION

With recent trends of low-cost sensors, miniaturization of computing and wireless communication, it is easier than ever before to obtain real-time data about indoor occupancy. There is a significant interest in smart services for indoor environments as organizations aim to improve energy savings, improve facility utilization (Trivedi and Badarla, 2019), and want to implement social distancing in the context of the recent covid19 pandemic. Due to the pandemic there is an increased interest in crowd control and the smart guidance of people in order to prevent that places get crowded. Smart sensor-technology could help showing people where a lot of room is to make it easier to keep distance from each other. However, for organizations it is still a challenge to obtain sensor data and to develop a model to gain accurate occupancy estimations, mostly due to hardware limitations and cost. While commercial solutions with accuracy levels down to a centimetre are available, they offer a poor business case for most organizations' purposes. At the same time, organizations store readily-available, but unused data, which could be used for this purpose, such as schedules, WiFi logs, and employee/client log-in data. We argue that combining this data with low-cost sensors can also yield a level of accuracy that will match most business needs. In light of this, we suggest a data aggregator information service combining available data with low-cost sensor data to provide an adequate occupancy detection accuracy at the lowest cost. The Smart Campus of the University of Twente (UT) in the Netherlands will be used as a case for this study. The UT is already estimating people count numbers using WiFi Access Points, but not many insights nor applications are based on that. The facility management department is interested in further enrichment of the data in order to provide better insights about the traffic of students and levels of space utilization. At the same time, the data about building-occupation can also be provided to students in order to facilitate their spreading around public study spaces, by recommending the least crowded buildings. This smart routing became high-priority for many campuses given the recent outbreak of the coronavirus and the social distancing measures that followed.

The prototype of our occupation information system is designed for application in a university campus environment. The University of Twente, which serves as a testing environment, includes eleven buildings that house lecture rooms, bookable study rooms and public study spaces available to students. The campus-wide WiFi network is used to estimate lecture attendance and to create occupation heatmaps. A recent project added Passive Infrared (PIR) sensors to bookable study rooms in the library to see if booked rooms are actually used. However, data about public study spaces is still lacking and this could be of value for finding quiet study places and to warrant social distancing in the near future due to the corona virus risk. The main aim of this study is to design a service that could inform students about where space is available to study based on sensor-data.

The paper is structured as follows. First, the methodology is described which explains the three design science cycles in this paper. Second, a semi-structured literature review is carried out to gain better understanding of the current developments of sensors. Expert interviews provide additional feedback on sensors and different aspects of designing a guidance service. Afterwards, the case is presented with the University of Twente as a testbed. Finally, we present our discussion chapter including limitations and a reflection on the value of the findings.

2. METHODOLOGY

The methodology of this study is based upon the design science research approach (Hevner and Chatterjee, 2010), which consists of three different design science research cycles: the relevance cycle, the rigor cycle, and the design science cycle. This chapter describes the three research design science cycles of Hevner and explains how this study complies to the design science approach.

The rigor cycle discusses the already existing knowledge to the research project (Hevner, 2007). This cycle checks artifacts and applications found in the application domain, already existing theories and methods, and expertise and experience on the artifact. The rigor cycle has been applied to this study by performing a literature study on the meta artifacts, existing theories and methods. Afterwards, interviews have been conducted to obtain knowledge in the form of expertise and practical experience. Experts in the field of the research domain have been interviewed: IT, facility management and the University of Twente.

The relevance cycle describes the application domain and the problems together with the opportunities (Hevner, 2007). The problems and opportunities of the research are described in the introduction of the paper. People, organizational systems, and technical systems are part of the application domain and together describe how the newly designed artifact will be implemented in the already existing environment. Therefore, the stakeholders of the artifact are discussed in this chapter. In addition, to properly implement the artifact several user, functional, and quality requirements have been established.

The design cycle generates design alternatives and evaluates these designs against the requirements until a satisfactory design has been achieved (Hevner, 2007). Based on the correspondence of the sensors obtained during the rigor cycle and the requirements established in the relevance cycle, decisions were made on which sensor to use in the design phase. In addition, the prototype has been established to fulfil the requirements in several design iterations.

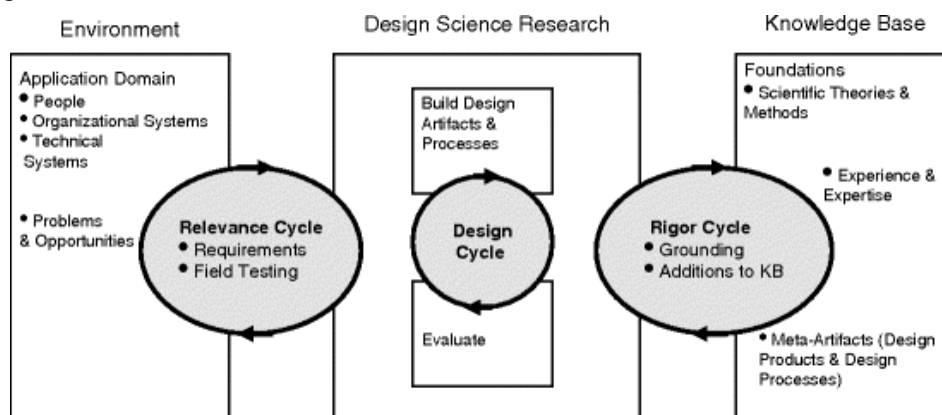


Figure 1. Overall view of the design science research cycle of Hevner (Hevner 2007)

3. THEORETICAL BACKGROUND

We carried out a literature review here with articles from the scientific database Scopus to inspect the current state of sensor research. The search terms “indoor localization” and “occupancy detection” were used to find trends and general information about similar systems. A list of sensor technologies was derived from the initial search that were used subsequently as keywords to find more detailed information for each type of sensor. Below is a summary for each sensor type identified, followed by an overview table.

3.1 WiFi

WiFi Received Signal Strength (RSS) has been a highly attractive method for indoor localization due to the wide deployment of WiFi Access Points (AP) and the fact that nearly no extra technical costs are involved (Guo *et al.*, 2020). However, RSS accuracy is notably low compared to other solutions, and is furthermore very sensitive to changes in dynamic environments. Any object can scatter the WiFi signal in unexpected ways, thus even a small change in the observed indoor space can cause big changes to the received signal, making it unreliable on its own. There are, however, efforts into increasing the accuracy by combining different techniques, such as machine-learning (Guo *et al.*, 2020). Furthermore, the ubiquity of WiFi APs and their ease of use with little extra cost allows this method to be used as a complement to enhance the accuracy of such a system.

3.2 Bluetooth Low Energy

Bluetooth Low Energy (BLE) is a wireless personal area network technology created by the Bluetooth Special Interest Group. This technology has been widely adopted among mobile device manufacturers, where it enables phones to communicate with small task-specific devices, such as Bluetooth headsets. Compared to ordinary Bluetooth, BLE requires considerably less power and is optimized for transferring small data packages (Tekler *et al.*, 2020). Common deployments of BLE for indoor positioning require a number of beacons or tags to be placed in the location of interest. The beacons broadcast a small data package to nearby bluetooth-enabled devices (readers). This is currently used in shopping malls for pushing advertisements to nearby users. Since BLE beacons are one-way communication devices, they can't return information about nearby devices (Borić, Díaz Redondo and Fernández Vilas, 2018). Instead the device uses the received data to estimate its location or sends it to a server. Different locations within the environment will have a different set of RSS, called a fingerprint. The reader collects RSS data from all nearby tags and sends this data to a server. The server has a list of predefined fingerprints, corresponding to specific locations with the associated RSSI received at that spot. The k-nearest neighbour algorithm is often used to establish at which fingerprint the user is located (Tekler *et al.*, 2020). Another method is to have the device estimate its position based on the signals received from beacons, however this is commonly applied for indoor localization systems instead. Researchers state the low price and power consumption as well as the ubiquitous adoption in mobile devices as the main advantages of BLE (Filippoupolitis, Oliff and Loukas, 2016). The user is, however, required to use an app for this method to work.

3.3 Passive Infrared Sensors

Passive Infrared Sensors (PIR) can be used to detect a movement of an individual and to detect the direction of movement. It is able to sense direction because it contains multiple sensing areas which combined output can be used to detect direction. The accuracy of PIR by itself is not always very good. It has the problem of cumulative errors since it cannot correct itself over time. Using the PIR by itself only gives an accuracy of around 27% for estimating room occupation over time. However, this can easily be improved to 90% when combined with other door sensors (Gaonkar *et al.*, 2019). One major drawback of PIR is that it only works in areas with clear doors. Since it works like a door counter it is not possible to deploy it in open areas without clear entrances.

3.4 Video Camera Recognition

Image recognition systems based on video cameras are currently used to determine the availability of parking spaces in car garages and cities. Cameras are used together with an IoT device to monitor the status of a parking spot (Ling *et al.*, 2017). In addition, this system was able to track the shapes and sizes of vehicles (Ling *et al.*, 2017) based on real-time processing of the data. These findings are useful in determining the occupation of seats in public places as well, since a camera can be used to check if a seat is occupied by a person or by, for instance, a bag. Machine learning can be applied to improve the recognition. As a result, the accuracy of using the image detection system is high according to completed experiments (Ling *et al.*, 2017). Hence, using cameras together with an image detection system would be suitable for the research of seat occupancy. Furthermore, real-time face-detection systems are used to identify and authenticate people in a specific picture of video. Statistical pattern recognition methods are used to recognize faces (Kollreider *et al.*, 2007). In addition, algorithms used to determine facial recognition exist, for instance, the Eigenfaces technique building onto the principal component analysis, the Fisherfaces algorithm, and the hidden Markov model (Mitchell, 2010). Currently, advances in technology make it possible to adjust the existing algorithms to achieve significant accuracy like 92% in the study of Pillay (Pillay and Viriri, 2019). Thermal imaging has also been applied to more easily and accurately distinguish between humans and non-living objects.

3.5 Carried Sensor

To locate a person, it is also possible to provide this person with a sensor, like an RFID tag which users use to send their location. This concept has already been tested with cars (Mainetti *et al.*, 2014). It is also widely adopted in warehouses for tracking the location of goods. But as there is no way for the tag to distinguish between a free and an occupied seat, the only way to use them for occupancy detection is for everyone to wear a tag. Such systems for tracking humans are already used in health monitoring and some localization systems (Trivedi and Badarla, 2019). RFID tags give a high accuracy but this means that users need to wear the tag all the time and have many readers installed, due to their relatively low range. There should also be a thorough debate about the ethics of providing users with this technology. Since the location is determined inside the building this would have to rely on other sensors besides GPS. GPS does not work indoors. Barometers and compass data or acoustic localization can be used as additional positioning-sensors (Retscher, 2006; Liu *et al.*, 2020).

3.6 Ultra Wide Band (UWB)

UWB is one of the most recent, accurate, and promising technologies for indoor localization (Alarifi *et al.*, 2016). For UWB, Beacons are deployed, similarly to BLE together with some carried receiver. Common methods for positioning are calculating the angle of arrival, time of arrival or time difference of arrival of the signal. Due to having very high bandwidth (100 Mbps) and extremely short wavelength of the signal it offers a high-accuracy positioning, even in the presence of severe multipath (Alarifi *et al.*, 2016). The signals effectively pass through walls, equipment, and any other obstacles. Errors can be brought down to the sub-centimeter range, making it suitable for systems with high accuracy requirements (Alarifi *et al.*, 2016). The main drawback of this technology is the high cost of the sensors and their infrastructure.

3.7 Pressure Sensors

Pressure sensors can measure the force applied to them for detecting an object. They have commonly been used in cars to detect the occupancy of passenger seats to assess if the airbag should deploy (Baur, Huber and Zeeb, 2007). In the case of buildings, the sensors can also be directly built into chairs. This method provides excellent accuracy for a low cost per sensor. However, it cannot distinguish between a student and his bag. A more expensive pressure sensor that uses electric field sensing can be used to achieve this (Tumpold and Satz, 2009). Capacitive pressure sensors sense the difference impedance between two areas of a chair to detect the presence of a person. It has the benefit of being contactless so it can be placed unobtrusively.

Despite its higher cost compared to traditional pressure sensors, it has no wearing parts, which reduces maintenance costs. They provide both rich and accurate tactile data but need to be installed in the seat.

3.8 Acoustic Sensors

Another method which can be used for occupation detection is acoustic sensing (Trivedi and Badarla, 2019). This method consists of a microphone that listens for any sound and optionally speakers that play a constant sound at the border of human hearing to be picked up by the microphone. It is, however, not very suitable for occupation estimations beyond detecting presence, and even then it is rarely used on its own, due to common false-positives and false-negatives (Trivedi and Badarla, 2019). In noisy environments when there is no direct path to the target it is a challenge to select a frequency of the sound for such a system (Liu *et al.*, 2020).

4. EXPERT EVALUATION

In one of the first relevance cycles of the design science approach, we conducted a series of face-to-face semi-structured expert interviews. Table 1 displays the list of experts included in this evaluation. The interviews mostly confirmed our findings from literature, but also provided additional insights into various aspects of the information service design. The most important points have been summarized per category.

Table 1. Expert selection including purpose and background

Interviewee	Purpose	Background
Interviewee A	IoT-Sensor-technology reliability check	Principal consultant at a smart IoT technology company
Interviewee B	IT-infrastructure possibilities overview	Director of a semi-public IT infrastructure company
Interviewee C	University information services control	Information service specialist based on smart WiFi networks
Interviewee D	Potential use and business purposes	Data-driven facility services planning expert

The choice of source data is not only dependent on the organization's budget and required accuracy, but also the type of environment one wants to detect occupation in. For example, fixed static environments allow accurate results using Time-of-Flight or RSS-based technologies, due to low variation of the signal. But if the area has a fixed entrance, it would be much easier and cheaper to use a simple PIR sensor instead. (Interviewee A). The non-sensor data available is also context dependent. In our case study, a roster schedule is available for the lecture rooms around public study spaces as well as booking data about study rooms, giving us an expected number of students that need to be deducted from the total count. The company from Interviewee D is a facility cleaning service supplier, for example, utilizes flight information from airports. Not all non-sensor data can be useful in this regard, though. Interviewee D noted that feedback systems are not reliable in real-time or short-term people counting, as the percentage of people who give feedback fluctuates too much.

4.1 Sensor-Data Review

Device-reliant solutions such as WiFi, Bluetooth and mobile sensors may not accurately account for the number of people. According to Interviewee B, disadvantages of sensing devices, instead of humans are that several devices can be worn by the same person, or the device might be set up to not be discoverable by the beacons or access points, or the device measuring in a different area as expected, all of which skew the results.

Two general types of approaches using mobile devices have been identified both from literature (Trivedi and Badarla, 2019) and Interviewee C: device-based and network-based. Network-based approaches are characterized by having an infrastructure of beacons or access points that can listen for nearby devices. When using such an infrastructure, there is no need for the users to install an app, making it more flexible (Interviewee C). On the other hand, you can't count devices that have their WiFi/Bluetooth turned off. This approach can give a good estimation of the number of people nearby. However, it's possible to connect to an AP from another room and deduplication in case of multiple devices is not always accurate. As such, results lack room-level granularity, but can be accurate on floor-level. Device-based approaches require the user to install an app on their mobile device. This method has a higher accuracy compared to network-based WiFi and BLE, due to using additional data sources from the device's sensors to pinpoint the location. In this case the device listens for nearby beacons. The collected data is either used to estimate location using a local database or the information is sent to a server, and the device receives its location back from that server. This approach is more commonly used for localization and navigation systems, however, as you can only see the devices that are actively using the app to navigate at this moment. In light of this, real-time data can't be accurate without a very high adoption level of the app.

Passive infrared sensors are already widely employed in office buildings for controlling automatic lights, blinds, heating and many others. Because of its cheap price and low power consumption, PIR is recommended in solutions where low level of detail is sufficient. For accurate people counting only the use of PIR sensors is not enough. There are also PIR door counters which are quite accurate short-term (95%+, Interviewee D), but they suffer from having an accumulating error over time (Interviewee C). For example, two people walking hand-in-hand are likely to be counted as only one person.

Image recognition is another popular method in practice. Advances in machine learning and AI have led to algorithms that can even recognize faces and objects in less than a second without high computational requirements (Interviewee A). Thermal imaging improves accuracy in distinguishing humans even further. However, while experts consider cameras relatively cheap, such a solution is more costly compared to PIR or using already available WiFi infrastructure (Interviewee B). However, each expert noted that cameras add privacy concerns that need to be addressed. There are ways to circumvent this, for example by analyzing the image offline and sending only the people count and deleting the image (Interviewee B) and installing the camera overhead so that faces are not visible (Interviewee C).

Pressure sensors installed on seats are a very easy way to get accurate occupation information. Such a method is applied for example at a big Airport (Interviewee D). It is a good approach for seat-level accuracy for indoor locations with a relatively small number of fixed seats. The sensors themselves are very cheap, but the total cost of the system in our case exceeds most other options, due to the large quantity of sensors required. The University of Twente, for example, has 10,000 chairs that would need to be equipped with a sensor as well as with a battery and a networking device (Interviewee C).

4.2 Business Value

The experts shared the view that there is a growing demand for indoor occupancy data. Information about the level of utilization of buildings is needed for management decisions aimed at better space management. However, it's important to note the accuracy of the data when taking management decisions.

While there is growing demand for indoor occupation detection, there isn't much demand for high positioning accuracy. Instead organizations are interested in the cheapest solution that will provide a suitable level of accuracy according to their business needs. While a seat-level accuracy can provide better insights, the additional cost usually doesn't warrant the much higher investment required. However, according to interviewee B there is demand for high positioning accuracy in the fire department. The fire department wants to know if there are still people inside of a building and where these people exactly are to rescue them. Students and customers in general demand better information services, but also expect them for free. Furthermore, having an intrusive method of occupancy detection such as cameras is mistrusted, even if the goals of the camera placement are clearly stated by the organization.

4.3 Privacy Concerns

All experts noted the importance of ensuring privacy in such an information system. One of the recommendations was to minimize the personal data stored on the system. Images can be analysed by an AI counting people internally, sending only count estimates over a network, not storing any images. Furthermore, from the point of view of a provider of such an information service, it's not recommended to store data on own servers, but rather create a self-contained system at client's location, using clients' data servers. MAC addresses are also considered personal data, and as such should be anonymized. Interviewee C also notes that MAC addresses should not be traceable for longer than a day. While privacy control is very important, interviewee D noted that the system still should be designed in a way that allows sharing the data between the stakeholders interested in it. As a contractor of the UT, having access to the occupation data can help them provide a better service to the university. Table 2 gives a summary of the literature and expert's evaluations of sensor data technologies.

Table 2. Summary of the sensors identified from literature and interviews

Technology	Advantages	Disadvantages
Device-reliant		
WiFi	Uses existing infrastructure, unobtrusive	Low accuracy, negatively affected by dynamic environments. Device de-duplication challenges
BLE	Low power consumption, high adoption among mobile devices	Needs users to actively use an app, needs beacon infrastructure
Mobile sensors	Available in all modern devices. Combining all sensors provide high accuracy	Needs users to actively use an app and allow permissions to access sensor data
UWB	Very high accuracy positioning, even without direct line of sight.	High cost of UWB equipment, app use, challenge of UWB antenna placement
Self-sufficient		
PIR	Cheap, low energy consumption, proven and widely-used tech	Accumulating error, more useful for detecting presence
Camera	Very high accuracy, high granularity of occupation data	Social obtrusiveness and privacy issues
Acoustic	High location accuracy (2mm-1.2cm depending on setup), low time latency (6-250 ms)	Severe attenuation of aerial acoustic signals, efficiency reduces as distance increases, no counting
Mounted/wearable		
Pressure	Very accurate in 10-100kg range	Needs to be installed on each seat. Needs additional IoT components to send data over a network.
RFID	Cheap tags, low power consumption	Requires people to wear RFID tags, reader needs to be in close proximity to read the tag; tracks people/objects, not occupation

5. CASE STUDY FOR VALIDATION OF THE DESIGN

In this section, we relate our findings to the specifics of our case study. We begin the relevance cycle by describing the case of the University of Twente (UT). Requirements derived from the stakeholder analysis are applied to select the most appropriate scenario out of possibilities identified in the rigor cycle. Finally, we evaluate the feasibility of our idea by carrying out a short feasibility study.

Considering the Students' perspective, they desire to find a free seat in a public study place. They want the proposed information system to appoint them a building, floor of a building, or area of a floor which they are sure of there is an empty seat or a less crowded environment. The application for the students should have an overview of all the buildings and a simple indication system which points out the relative occupancy of

seats in public study places. To make the application more friendly and easy-to-use, a filter and recommendation option should be added.

The second stakeholder in the IS design is the university itself. The facility management department do not have proper information on the occupancy of seats in public study areas. They need this information to determine if there is a need for additional seats in public study areas. Having this data may result in a better utilization of space overall. So, for the university the IS should show a clear overview of the occupancy rates of each building, floor of a building, area of a floor, and in total. Moreover, the users should be able to filter on the building, floor of a building, and area of a floor to give them a more general overview.

The final stakeholder group consist of third parties such as the cleaning service. For third parties it is valuable to obtain the location which currently has the lowest occupancy, so they are able to clean a place without much interference of students studying in a certain place. This IS will give them additional information about public study places. Cleaning services should be able to see the relative occupancy rate of each building, floor of a building, and area of a floor to determine the least busy public study places. Interviewee D argued that the cost optimization is mostly a matter of perception.

Table 3 gives an overview of scores of the different sensors which could be applicable at the UT in public study places together with the quality requirements specified. WiFi is chosen as the baseline and therefore all other sensors are compared to the scores of using just WiFi. A scale of the three levels +, 0 and - has been chosen to keep the overview as clear as possible and to determine the total score of the sensor(s). These levels have the following descriptions: + shows that a requirement is fully met and is thus a positive level, 0 displays that a requirement is partially met and is thus between the positive and negative level, and - describes that a requirement is not met and receives a negative level.

Table 3. An overview of different sensor combinations with their scores on the quality requirements

Sensor type/Quality requirement	Accuracy		Ease of implementation		Legal concerns	Social obtrusiveness	Implementation costs		Total score of the sensors
	<i>Number of people accurately measured</i>	<i>Positioning of people accurately measured</i>	<i>Low amount of effort to create a system</i>	<i>Low amount of effort to obtain devices/equipment</i>	<i>Low data privacy concerns</i>	<i>Low social obtrusiveness of the sensor</i>	<i>Low costs of implementing the system</i>	<i>Low costs of obtaining devices/sensors/equipment</i>	
WiFi (baseline)	-	-	+	+	0	+	+	+	+++
WiFi + BLE	0	+	-	-	0	0	0	-	--
WiFi + Infrared	0	-	+	0	0	0	+	0	+
Cameras	+	+	0	0	-	-	0	0	0

The first scenario considered is using just WiFi to determine the seat occupancy in public study areas. As can be seen in table 3 a significant amount of the requirements is met when WiFi is used. The ease of implementation requirement is met, since the UT already has a system in place and uses it to determine the number of people on a certain floor in a certain building on the UT. In addition, the social obtrusiveness is low, as students' phones are used as a sensor without them knowing this happens. Furthermore, data privacy concerns are low, since the data obtained is only used to create new value for the students and does not harm them in any way. Moreover, the data is not publicly accessible, but only when connected to the WiFi of the UT. As a result of the infrastructure being in place, the costs will be low. Therefore, the cost requirement is met as well. The only downside of using WiFi is the low accuracy. As a result, only a rough estimation of the number of people and their position can be given.

An alternative to the scenario just using WiFi is combining it with BLE. The main advantage of this scenario is the improvement of the accuracy. However, this scenario brings several disadvantages as well.

First, there is no existing system at the UT using BLE. In addition, there are no beacons to track the BLE signals from the mobile phones of students. As a result, the costs of creating a system will be a bit costly and a significant number of Bluetooth beacons have to be bought resulting in a high cost of devices. Third, the accuracy of BLE is high when using the right version, but when using an older version the accuracy depends on if you have your Bluetooth turned on or not. Fourth, the Bluetooth beacons have to be installed and are visible to students which result in a small social obtrusiveness. In conclusion, combining WiFi with BLE will result in a small increase of the accuracy, but add many disadvantages which do not occur when using only WiFi.

The third scenario consists of combining WiFi with infrared sensors. Adding infrared sensors will increase the accuracy of the number of people in a certain study place. However, an infrared sensor might detect the bag on your back as another person. This error cannot be filtered out during the day and will slowly result in a bigger error in the estimation of the number of people. An additional downside of infrared sensors is the utility. Infrared sensors can only be used in public study places where the entries and exits are clear and not in large public areas, since you cannot determine where people are in large areas using infrared sensors. Another small disadvantage is the fact that infrared sensors have to be installed in some areas which students can notice. One major advantage of using infrared sensors is availability in certain buildings of the UT. In addition, the data privacy concerns are low, as only the number of people in a certain public study place are saved. In conclusion, infrared sensors can improve the accuracy without having to deal with additional disadvantages.

The fourth scenario considered uses only cameras. On first sight, using cameras to determine the occupancy of available seats in public places seems favourable, due to the high accuracy of cameras. However, there are a significant number of disadvantages at the ease of implementation, legal, social obtrusiveness and cost side. There are already cameras being used at the UT, but not connected to a system which measures the number of available seats in public study places. Therefore, the system has to be tweaked to work in this study. In addition, cameras have to be bought to install one in every public place. This will result in a social obtrusiveness concern as well, since students are being filmed. The data privacy concerns and social obtrusiveness concerns have to be taken into account when choosing cameras. Since the infrastructure is not fully in place, additional costs have to be made as well. In conclusion, cameras are highly accurate, but add a significant number of disadvantages.

After exploring all four scenarios, we decided to choose the first scenario using WiFi only. According to the literature and experts interviews, this scenario fits the project best. As can be seen in Table 3, WiFi meets the ease of implementation and costs requirements. However, the accuracy requirement is not totally met. As a result, we keep the low accuracy in mind when designing the prototype by not using the exact number and positioning of people. Instead it will give a rough estimation of the number of people and an indication of the positioning of people per floor instead of per table in a public study room.

Figure 2 shows the Archimate 3.1 diagram of a layered high-level architecture of our proposed service. Certain elements are combined, simplified or removed from the model to preserve clarity. Starting from the technology layer, data recorded from WiFi access points and PIR sensors (currently only in the library) is stored on the UT databases, along with Roster data for lecture rooms and Booking data for study rooms. Anonymized data from these sources, limited to just people count, location and time will be drawn into the backend of our app. The backend is also hosted on the UT servers. Aside from the data, the backend also contains two application components. The Occupancy analysis and prediction which takes into account both real-time and historical data. It uses a regression model to predict future occupation. The recommender algorithm uses near real-time data to suggest a ranking of least occupied spaces fitting customer criteria. The backend is accessible via two separate endpoints.

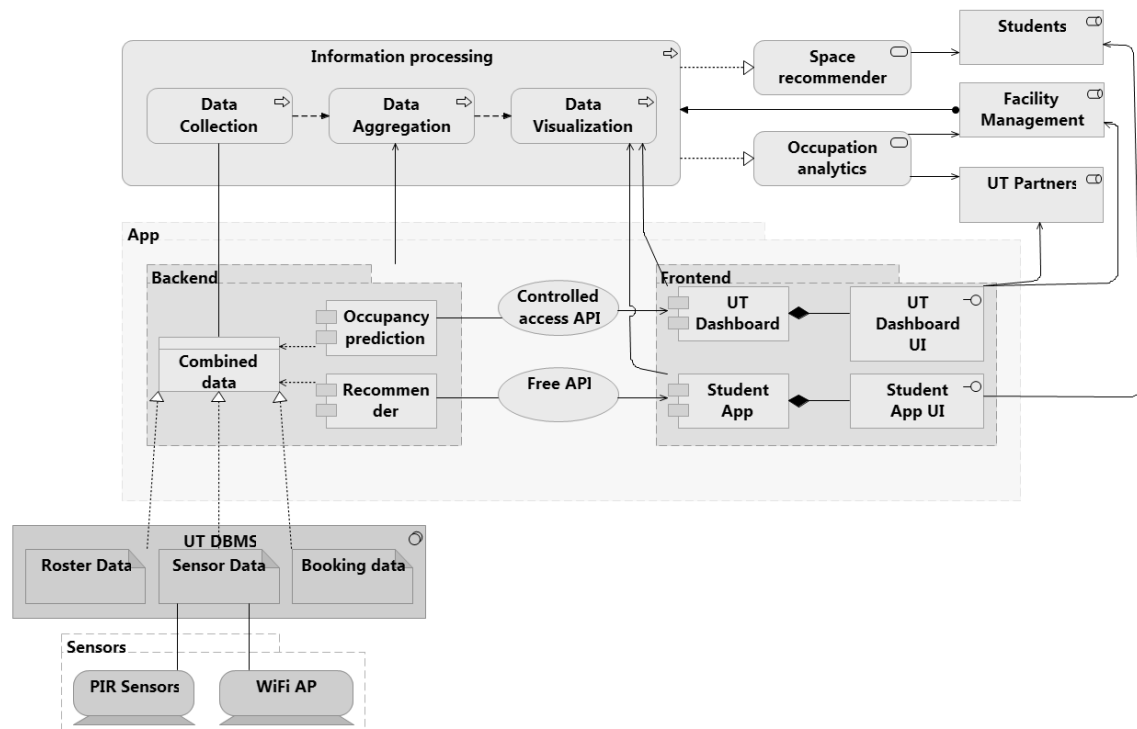


Figure 2. An architecture of an occupation guidance service

The recommender is available to the Student app via a non-restricted API. Historical data from the analytics and prediction component, however, is more sensitive and will only be accessible via a secured login on the UT dashboard page. The backend of the application layer takes care of collecting and aggregating the data, while the front end serves as means of visualization. The processes of collecting, aggregating and visualizing the data realize the two services our system is offering to students and the UT. Third-party contractors may also be allowed access to the dashboard at the UT's discretion. Responsibility for managing the process is assigned to the Facility management.

The occupation information service will be displayed in two separate apps - one for students and one for university staff and they will rely on data already available in the UT's database (WiFi and roster data). The student app will contain a recommender algorithm and a filter to help students find public study spaces faster and easier based on their preferences. The UT app will be a dashboard showing occupation data that can also be filtered based on the user's preferences.

We try to use as much existing infrastructure as possible. Hundreds of WiFi APs are available on campus. Our other stream of input data is the rostering schedule for the rooms around public study spaces. Server and database costs will also be negligible as data is already stored on UT servers and we assume only a minimal amount of additional overhead will be incurred from incorporating our system. The apps that need to be developed will use existing design from the UT app, are not complicated and will be cheap to develop. The system will also have an API developed for it that will facilitate controlled access to the UT databases.

6. DISCUSSION

We reviewed literature on occupation detection systems to gain insights about the current state of the field. Our focus was on the various sensor technologies being used for data generation and the overall design of such a system. What we found was that complete designs for an occupation detection system are hard to find in scientific literature, however each separate sensor was extensively researched. Most papers concerned with design were focused specifically on the algorithms used to extract information from the sensor data to improve the system's accuracy. However, despite the focus on accuracy, it is also difficult to obtain a good

understanding of the accuracy one can expect from each sensor. The reported accuracy numbers vary a lot across papers due to different methods and assumptions, and it is hard to generalize from those results. Also, most papers focused on only one data source. Few papers analyse the potential of combining sensors and the benefits of doing so are still not very clear.

Most papers acknowledge the potential insights that can be gained by analysing occupation data such as “energy savings, user comfort and degree of automation” (Trivedi and Badarla, 2019). However, estimating the exact extent of those benefits seems elusive. Our expert interviews also confirmed that companies are interested in occupation insights. However, they noted that businesses don’t want to spend a lot of money on this, as justifying the expenditure is difficult. We also identified several gaps between literature and practice. The high accuracy reported in literature doesn’t match the estimates of our experts. It remains up for debate whether accuracy is overstated in scientific literature, or if practitioners are not making the most of recent advancements in the field.

We argue that the perceived high cost of the investment, combined with uncertainty in the potential benefits are the main factors preventing faster adoption in a business setting. Research that quantifies potential benefits is needed in our opinion. This is, however, difficult, as the design and therefore cost of such a system depend highly on the business goal and setting. All experts also mentioned data privacy as a challenge for this type of systems.

Future work should expand the occupation information service’s information processing. Currently a basic recommender and a general descriptive analytics dashboard are included for simplicity’s sake. Designing a recommender that takes into account location proximity would be an important addition for the students for example. Furthermore, user accounts and social contact information can be included in the information processing, allowing user interaction and richer data. Those additions are expensive, however, and require further cost-benefit analysis.

The previous limitation is partly due to the COVID-19 (coronavirus) outbreak which happened in the middle of the study. The coronavirus led to a closure of the whole University of Twente campus, which made the existing infrastructure measuring the number of people in an area based on WiFi useless, as there were no students in the University of Twente buildings. As a result, this is a significant limitation to the study.

Another limitation to the study is the decision to use one central scenario in the designing phase. In this case, the WiFi as a baseline is chosen in the design. However, creating and testing other sensors or combinations of sensors based on data triangulation could have led to different insights which are not explored. It is suggested to, in future work, explore different scenarios where other sensors and combinations of sensors are explored in the same environment. For instance, looking into PIR sensors which the University of Twente uses to estimate the number of people in a specific area as well. Also, we did not include the combination with existing telecom signals such as 5g which are also more convenient using in a outdoor situation.

In this paper we presented a preliminary design of an occupation information service - a potential new niche in the information market but potentially very useful in a social-distancing guidance application. While there seems to be a growing demand for such a service, there are several challenges that a viable business model needs to overcome - cost, unclear benefits and data privacy concerns. We proposed a low-budget solution, showing that the cost of such a can be minimized by a combination of available infrastructure and data sources. We reviewed different source systems for gathering occupation data with their advantages and disadvantages with regards to business requirements. The implications for business are that when high accuracy of occupation data is not necessary, estimations made by matching already available data can fulfil needs at a low cost. The business needs to approach occupation information systems with a clear idea of how to use the information and its value, however, as the benefits are still difficult to quantify or generalize.

REFERENCES

- Alarifí, A. *et al.* (2016) ‘Ultra wideband indoor positioning technologies: Analysis and recent advances’, *Sensors (Switzerland)*, 16(5), pp. 1–36. doi: 10.3390/s16050707.
- Baur, R., Huber, H.-P. and Zeeb, E. (2007) ‘(12) United States Patent’. United States. Available at: <https://patentimages.storage.googleapis.com/c1/c6/af/f51bc6e04e48e6/US7162111.pdf>.
- Borić, M., Díaz Redondo, R. P. and Fernández Vilas, A. (2018) ‘Space Occupancy through BLE Dynamic Broadcasting’, *Wireless Communications and Mobile Computing*, 2018, pp. 1–17. doi: 10.1155/2018/2182614.

- Filippoupolitis, A., Oliff, W. and Loukas, G. (2016) 'Bluetooth Low Energy Based Occupancy Detection for Emergency Management', *Proceedings - 2016 15th International Conference on Ubiquitous Computing and Communications and 2016 8th International Symposium on Cyberspace and Security, IUCC-CSS 2016*, pp. 31–38. doi: 10.1109/IUCC-CSS.2016.013.
- Gaonkar, P. *et al.* (2019) 'Occupancy Estimation in Semi-Public Spaces using Sensor Fusion and Context Awareness', *Proceedings of 2019 IEEE Region 10 Symposium, TENSYP 2019*, 7, pp. 131–136. doi: 10.1109/TENSYP46218.2019.8971134.
- Guo, X. *et al.* (2020) 'Robust WiFi Localization by Fusing Derivative Fingerprints of RSS and Multiple Classifiers', *IEEE Transactions on Industrial Informatics*, 16(5), pp. 3177–3186. doi: 10.1109/TII.2019.2910664.
- Hevner, A. and Chatterjee, S. (2010) 'Design Science Research in Information Systems', in *Design Research in Information Systems: Theory and Practice*. Boston, MA: Springer US, pp. 9–22. doi: 10.1007/978-1-4419-5653-8_2.
- Hevner, A. R. (2007) 'A Three Cycle View of Design Science Research', *Scandinavian Journal of Information Systems*, 19(2), pp. 87–92. doi: <http://aisel.aisnet.org/sjis/vol19/iss2/4>.
- Kollreider, K. *et al.* (2007) 'Real-time face detection and motion analysis with application in "liveness" assessment', *IEEE Transactions on Information Forensics and Security*, 2(3), pp. 548–558. doi: 10.1109/TIFS.2007.902037.
- Ling, X. *et al.* (2017) 'Identifying parking spaces & detecting occupancy using vision-based IoT devices', *GIoTS 2017 - Global Internet of Things Summit, Proceedings*. doi: 10.1109/GIOTS.2017.8016227.
- Liu, M. *et al.* (2020) 'Indoor acoustic localization: a survey', *Human-centric Computing and Information Sciences*. Springer Berlin Heidelberg, 10(1). doi: 10.1186/s13673-019-0207-4.
- Mainetti, L. *et al.* (2014) 'Integration of RFID and WSN technologies in a Smart Parking System', *2014 22nd International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2014*. FESB, University of Split, pp. 104–110. doi: 10.1109/SOFTCOM.2014.7039099.
- Mitchell, C. (Albert N. S. of E. (2010) *Applications of Convolutional Neural Networks to Facial Detection and Recognition for Augmented Reality and Wearable Computing*. Albert Nerken School of Engineering.
- Pillay, P. and Viriri, S. (2019) 'Foresight: Real time facial detection and recognition using webassembly and localized deep neural networks', *2019 Conference on Information Communications Technology and Society, ICTAS 2019*. IEEE, pp. 1–6. doi: 10.1109/ICTAS.2019.8703634.
- Retscher, G. (2006) 'Location determination in indoor environments for Pedestrian navigation', *Record - IEEE PLANS, Position Location and Navigation Symposium*. IEEE, 2006, pp. 547–555. doi: 10.1109/PLANS.2006.1650643.
- Tekler, Z. D. *et al.* (2020) 'A scalable Bluetooth Low Energy approach to identify occupancy patterns and profiles in office spaces', *Building and Environment*. Elsevier Ltd, 171, p. 106681. doi: 10.1016/j.buildenv.2020.106681.
- Trivedi, D. and Badarla, V. (2019) 'Occupancy detection systems for indoor environments: A survey of approaches and methods', *Indoor and Built Environment*, 0(0), pp. 1–17. doi: 10.1177/1420326X19875621.
- Tumpold, D. and Satz, A. (2009) 'Contactless seat occupation detection system based on electric field sensing', *IECON Proceedings (Industrial Electronics Conference)*, pp. 1823–1828. doi: 10.1109/IECON.2009.5414836.