UNIVERSITY OF AMSTERDAM

FINAL REPORT TWEEDEJAARSPROJECT

Occupancy counts based on motion triggers

June 30, 2024

Students:
Mark Broerse
14651076

Joep Deeben
14613069

Thijmen de Groen
14588757

Surja Sarker
14007444

Olivier Stam
14613573

Tutor:
Maas Hermes

Group:
AIdea

Course:
Tweedejaarsproject BSc KI

Contents

Abstract

This project is conducted in collaboration with the University Library of the Universiteit van Amsterdam and aims to develop an accurate estimation tool for study space occupancy. With the end goal being to provide several groups of people a useful tool to give insights on study space occupancy and help them use their time more effectively. In order to achieve this goal sensor data from different infra-red sensors within the University Library have been analysed and a model has been created to facilitate these demands. In addition to a model, recommendations for different sensors have been made and different approaches have been discussed to achieve the final goal.

1 Introduction

This project has been carried out in collaboration with the University Library (UB) of the Universiteit van Amsterdam (UvA). The UvA is spread over different locations throughout Amsterdam. Most of these locations offer study spaces, with the library being one of them. The University Library aims to be a place where students, researchers and staff alike can safely collect, select, store and manage information, both online and offline (University Library 2024). Furthermore, the UB offers multiple different types of study spaces, which consist of open study spaces and smaller collaboration rooms. These collaboration rooms require a booking in order use them. While the room bookings give the UB an indication if the rooms are occupied, there is no way of knowing whether the rooms are correctly utilized and what the exact occupancy of these and the open study spaces are.

1.1 Problem Description

By providing researchers, staff and students information about study space availability and usage patterns the UB wants to make their life easier. The primary goal the UB wants to achieve with this project is to provide an accurate estimation of study space occupancy in both the open study spaces and project rooms. There are some simple intuitive ways to do this, like having someone physically count the occupancy or installing occupancy detecting cameras, but these methods are either not realistic as they are too costly or too privacy-invasive. Therefore, the UB uses infra-red motion sensors which are cost-effective and non-invasive. For open study spaces, sensors are placed under the desk, while collaboration rooms utilize sensors placed on the ceiling. However, these sensors do not completely solve the problem as they are unable to directly count the amount of people in a room or at a desk. They can exclusively count movement within their respective detection zone. The final goal of this project is to solve this problem and to make it possible to accurately estimate the occupancy in a collaboration room or study space based on the amount of detected motions at those spaces.

1.2 Product Vision

To reach the final goal, two main steps are necessary. First of all, a model that can accurately estimate occupancy based on motion triggers needs be developed and trained. Then, using this model, an application needs be developed to visualize and summarize this occupancy. This application needs to meet three distinct groups of users demands.

1.2.1 Users

1. **Students**
   First of all, the application needs to be useful for students who want to study at the UB. It is essential for students to have real-time insights into study space availability. By providing this information, students can plan their study sessions more effectively as they can be sure there is available study space. Likewise, less time will be wasted on searching for study space and more time can be put into actually studying.

2. **Library Staff**
   The next relevant group for this tool is the UB staff. Having access to real-time occupancy rates will enable staff to efficiently guide students to available spaces. Besides helping the students out, it will also help the staff with making an optimal plan for their responsibilities like up-keeping and cleaning. This way, work can be distributed more efficiently within the staff.

3. **University Library**
   The final relevant group that can profit from the application is the UvA as an organization itself, with the UB in particular. Currently, there are not enough study spaces to facilitate everyone. For the UB it would be helpful to know if their study spaces are used as intended. The collaboration rooms in particular are known to rarely be filled at max

capacity. If some collaboration rooms meant for ten people are consistently underpopulated, they could be reconfigured into smaller rooms to better meet demand. On the other hand, if a larger study space is always used at max capacity, it might be better to add more desks and chairs to increase the max capacity. This tool will help them optimize study spaces when accurate estimations can be made.

By addressing the demands of these groups, the project aims to create a tool that can be used by the UB that enhances study experience and optimizes study space usage within the UB while simultaneously helping out staff and the organisation. Once proven effective, this tool can be expanded and implemented across other UvA locations.

2 Sensors

The University Library currently uses two distinct types of motion sensors to detect movement; the Locus Sensor, which is placed under the desk, and the Alta Sensor, which is placed on the ceiling of a collaboration room. These sensors are able to report a number of detected motions (triggers) to the system in a set time frame of five minute intervals. Furthermore, the sensors will stop reporting to the system after fifteen minutes have past without detecting any movement. This is done to prevent the sensor from flooding the system with unnecessary zero values. Both sensors are made by the company Officebooking 2024 and are both integrated within their IoT-Core platform. The sensors use proprietary AI technology to minimize false positives. The section below will delve into the various aspects of the ESCA Locus and Alta sensors, including its design, functionality, and further information.

2.1 Locus Sensor

The ESCA *Locus* Occupancy Sensor 1 is an infra-red occupancy measuring device designed to enhance workplace management by trying to provide accurate and real-time data on desk occupancy.

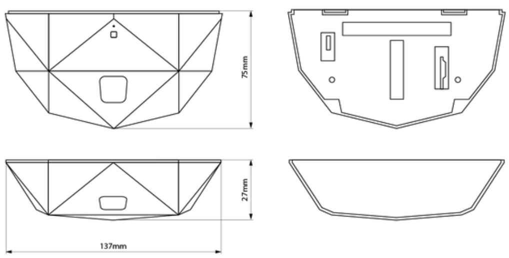

Figure 1: Diagram of the ESCA Locus Occupancy Sensor

The sensor is designed to be placed under desks in a non-invasive fashion, where it utilizes infrared light to detect occupancy. In this application it detects leg movement. The sensor's communication is based on the LoRaWan standard, which facilitates long-range, low-power wireless communication, making it suitable for large office environments, or in this case, study environments. This sensor has a one directional detection zone of about a meter in desk space. The received data within this project is of desks with a width of three meters with a singular Locus sensor placed in the middle underneath the desk. Therefore, further testing had to be conducted to figure out how useful the provided data was as this sensor is not inherently meant for this application, but rather for detecting whether someone is sitting at a single seat desk.

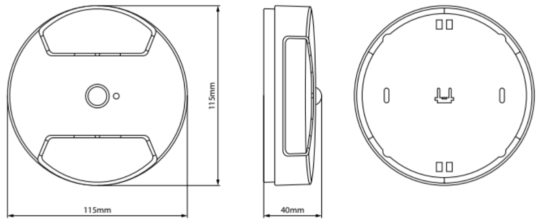

Figure 2: Diagram of the ESCA Alta Comfort and Occupancy sensor

2.2 Alta Sensor

The ESCA *Alta* Climate and occupancy sensor 2 is used for the collaboration rooms. Just like the Locus sensor, this sensor communicates using the LoRaWan standard. However, unlike the Locus sensor, this sensor is mounted on the ceiling of a collaboration room. This means that this sensor can pick up more general body movements, compared to exclusively detecting leg movements like the Locus sensor. It also has a larger detection range for picking up motion.

3 Method

This section will outline the research conducted during this project. It outlines the received data and some experiments that had to be conducted in order to examine the usability of the two types of sensors. Then, the model and its supporting algorithms, on which the final product is based, will be explained.

3.1 Data

3.1.1 Preprocessing

The provided dataset contained movement data of several study spaces (Locus sensors) and collaboration rooms (Alta sensors) throughout April and May of 2024. Each sensor had their own file. All data files had the same format, this format being: Date (with increments of five minute intervals), Motion counts (within such a five minute time-frame), and two columns that contained information on the connection strength of a sensor. These last two columns were deleted as this information was not useful for the data analysis. The pre-processed files ended up containing the two columns depicted in table 1.

Column name	**Dtype**
Date	datetime64[ns]
Count	float64

Table 1: Column Names and Data Types of Cleaned Data

Some entries in the data did not have a valid value for the motion counts, so these were deleted as well. Moreover, the data on the collaboration rooms was split into separate files for every month, so these files were merged based on the respective collaboration rooms. Furthermore, no target data was provided.

3.2 Experiments

In order to figure out if the provided data was useful and to obtain target data, several experiments have been conducted. Two experiments took place in order to do this for both sensors, these were carried out in the University library.

3.2.1 Locus Sensor Experiment

To test the Locus sensor a group of five people was seated in a *coupé*. A *coupé* consists of two, three meter wide desks, each desk having a maximum capacity of three people. Both these desks contain a singular Locus sensor mounted in the middle. The first part of the experiment attempted to discover how study habits influence the amount of detected movement. At certain times people changed their distance from the sensor in order to confirm its range and accuracy. In addition, some of the people left the *coupé* in order to see the impact this has on the detected movement.

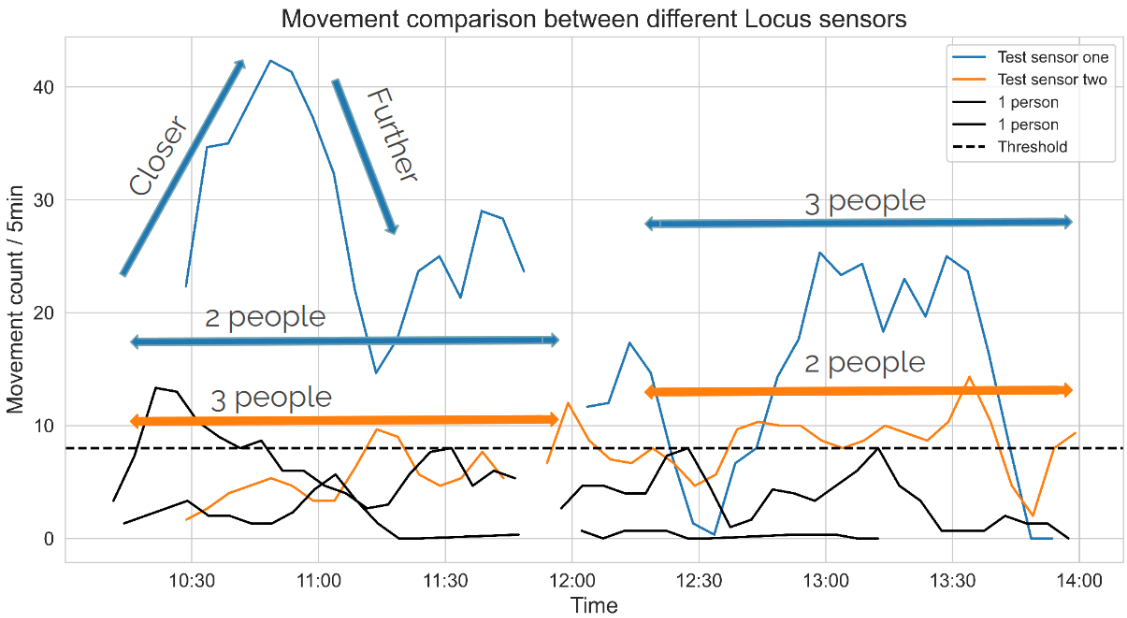

Figure 3: These are the first experiments results. Here different distances and amount of people have been tested.

As can be seen in figure 3, the distance from a student to a sensor greatly influences the amount of movement detected. Interestingly enough, there is a gap in the orange line. This occurred when the person in the middle stood up with the remaining two people sitting on both ends of the desk. This confirms the small detection zone of the Locus sensor. Another strange occurrence is that the same amount of people show up with different movement counts when comparing the first and second part of the figure. This proves that individual sensors might differ in accuracy, because of variance in the amount that students move whilst studying and how far they sit from the sensor.

3.2.2 Alta Sensor Experiment

The experiment to test the Alta sensor was conducted in a similar fashion. Four people were seated in a collaboration room meant for four people. In the particular room where the experiment was conducted, Locus sensors were installed under each desk in order to analyse those sensors further while testing the Alta sensor at the same time. Throughout this experiment, one extra person entered the room and two people left and re-entered the room several times. This produced target data for two to five people for the Alta sensor.

The Alta sensor's movement data seems more accurate, as figure 4 shows, indicated by the distinct lines between different occupancy levels. It is clearly visible when the fifth person entered the room. Later in the figure, around 12:30 a big fluctuation can be seen, an increase in movement as two people stood up and then a drop when they actually left. Indicating that there was potential to identify sessions and different occupancy levels within the data. The movement

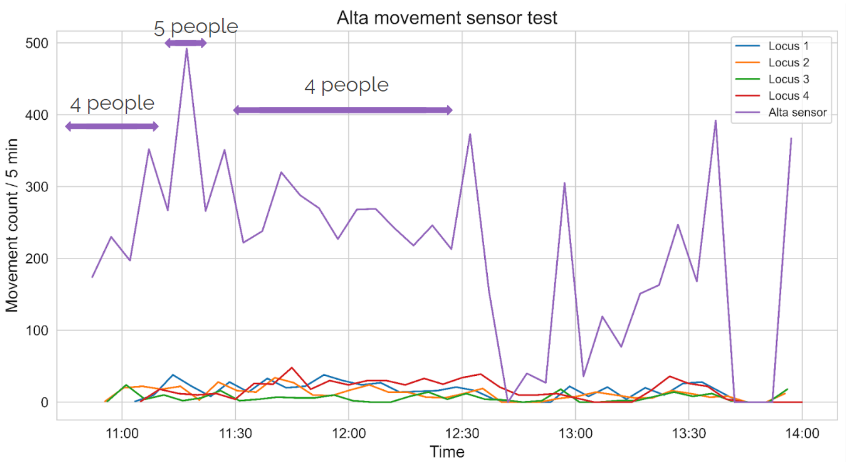

Figure 4: The results of the second experiment. Comparing different occupancy's within a collaboration room

counts of the individual Locus sensors can be seen in the bottom of the figure. They also show great variance within the detected movement, making our earlier findings more evident.

3.3 Solutions

3.3.1 Locus Sensor

As discussed, the biggest shortcoming of the Locus sensor is the detection range. On top of that, the difference in movement counts between different seating positions is significant. This was reported to our client by showing that the Locus can detect whether someone is occupying a single seat, but cannot be used in the application of estimating how many people occupy a three meter wide desk.

After visiting the sensor suppliers' website a possible solution was found. This is to change the Locus sensor to an *Omni sensor*. The Omni sensor 5 has a larger and 360° detection range. The sensor being from the same manufacturer, namely Officebooking, makes it easy to implement in the library's current sensor-infrastructure. This sensor will eliminate the short detection range shortcoming of the Locus sensor and therefore be a better fit for the large desks in the open study spaces. This new sensor might not succeed in estimating the exact occupancy, but will likely allow for a better estimation of whether the outer ends of the desk are occupied or not.

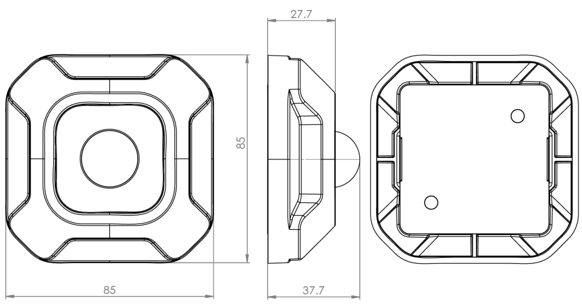

Figure 5: Diagram of the ESCA Omni Occupancy Sensor

3.3.2 Alta Sensor

The initial goal set by the product owner for the Alta sensor was to predict whether the project rooms are being used optimally. This involves determining if the number of occupants aligns appropriately with the room's capacity. After the test proved that it could provide accurate results, a model had to be made in order to estimate the occupancy. For this several steps had to be realized. Individual study sessions have to be detected. After this the average movement count of a session needs to be calculated without it being disturbed by outliers. Then finally, the occupancy estimation can be made.

3.3.3 Study Sessions

Using a model to predict occupancy per data point would lead to inaccurate results, a single point of variance in a students movement should not greatly impact the predicted occupancy in a room. To combat this the data needs to be segmented into study sessions which can be used separately to predict occupancy. A session is defined as a continuous period where movement is detected. The created algorithm starts by initializing variables to track session IDs and whether the data point is within a session. As it iterates through the data points, a new session begins when a non-zero count is encountered. The session continues until there are three consecutive entries with counts below a certain threshold (e.g., five), at which point the session is considered to have ended. This approach ensures that only significant periods of activity are classified as sessions. Sessions with fewer than four data points are discarded. This filtering step removes insignificant fluctuations that might otherwise be misclassified as sessions, ensuring that the identified sessions only represent actual study sessions instead of noise.
Additionally, within each session, the rolling average of movement counts is analyzed to detect significant changes. A rolling average is computed over a specified window, and if a substantial difference in average counts is detected, the session is split into new sessions. This approach ensures that periods with different levels of activity within a session are accurately identified and classified separately.

3.3.4 Average Calculations

When predicting the occupancy of project rooms an average of a study session needs to be calculated which can be put into our model. To prevent extreme values which are often found when students are entering or exiting the study space from skewing the session averages, an outlier removal process is implemented. The Interquartile Range (IQR) method is used to identify and remove outliers. Specifically, the first quartile (Q1) and the third quartile (Q3) of the data are calculated. Then the IQR is determined by subtracting Q1 from Q3. Data points that fall below Q1 minus 1.5 times the IQR or above Q3 plus 1.5 times the IQR are considered outliers and are excluded from the average calculation. By filtering out these outliers, the resulting average count for each session more accurately reflects typical movement levels. Once the outliers are removed, the average count for each session is calculated using the remaining data points.

3.3.5 Prediction Model

To estimate the occupancy within sessions, a relationship is established where the occupancy count is compared to the closest mean of the target data occupancy. The target data used is derived from the experiments and random manual occupancy checks. From this data the mean of the first standard deviation is taken as a robust indicator for occupancy, indicated by the red lines in figure 6. The movement count is then compared to these means where the closest one is taken. Resulting is an estimation of one, two, three, or four people.

However, the goal the product owner envisioned for the Alta sensor was to predict occupancy levels, specifically to determine whether there were more than three people in a given collaboration room based on movement data. To achieve this, a classification function was implemented. This function analyzes each identified session and classifies sessions with an estimated occupancy of one or two people as '1-2'. Sessions with an estimated occupancy of three or more people are classified as '3+'. This classification is simply made by adding an extra step to the earlier estimation classifying the two groups.

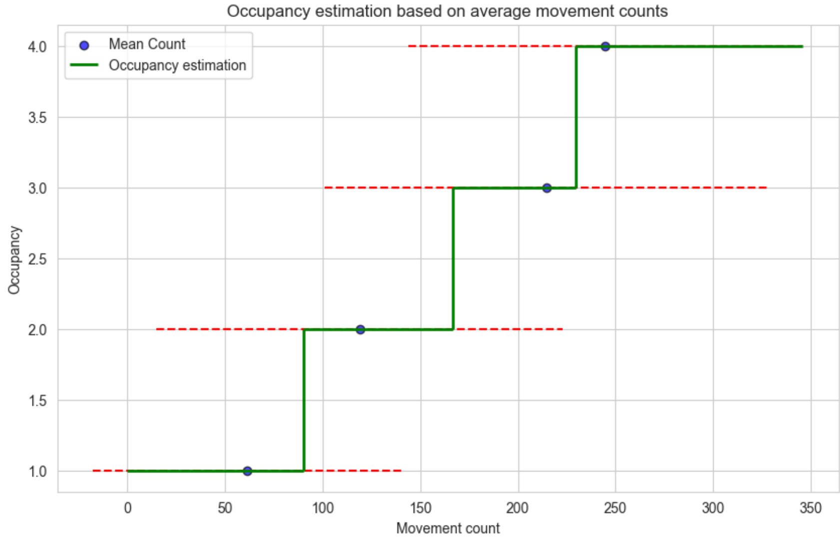

Figure 6: Occupancy classification, means displayed with 1std range

4 Product

The final deliverable was outfitted with a GUI to allow the client to easily utilize our research and models. Our preprocessing, algorithms and visualisations will be handled from the back-end so the user only needs to provide data in order to process it. The front end can be accessed through either a Jupyter notebook or through a web-based user interface which allows for non-technical library staff to interact with our deliverables. An option to choose between the exact occupancy or classified estimation was added in order to future prove the model. The exact estimation is inaccurate as of now as there is too little target data. When the client acquires more target data a simple additional file can be run to retrain the estimation model. If the target data is accurate enough, the exact occupancy estimation mode might be preferred.

4.1 Web-based UI

In the web-based user interface 7 the user is presented with an upload button where a raw data file needs to be uploaded. The model will handle the preprocessing so the raw data from the server can be used. Then the user can select the preferred time frame of the analysis, this can either be a single day or the entire time frame which is covered by the submitted data. After that the user can select whether they want to have the exact estimation which may lack some accuracy or the classified estimation which is more accurate but provides less detail.

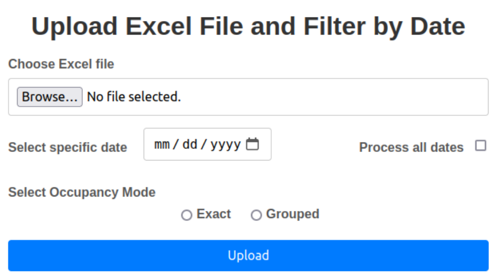

Figure 7: UI when launching the application

When all preferred options are selected and a file is uploaded, an estimation of the occupancy combined with a visualization of the split session of each selected day will be shown. This overview contains the duration, average amount of movements within a five minute interval, and estimated occupation of every session for every day within the selected time frame, as shown in figure 8. As an added request from the client, all the calculated data is gathered in a downloadable CSV file 2.

Column Name	Dtype
Date	daytime64[ns]
Session	Int
Total	float64
Mean	float64
Std	float64
Num_Points	float64
People	Int

Table 2: Column Names CSV file and Data Types

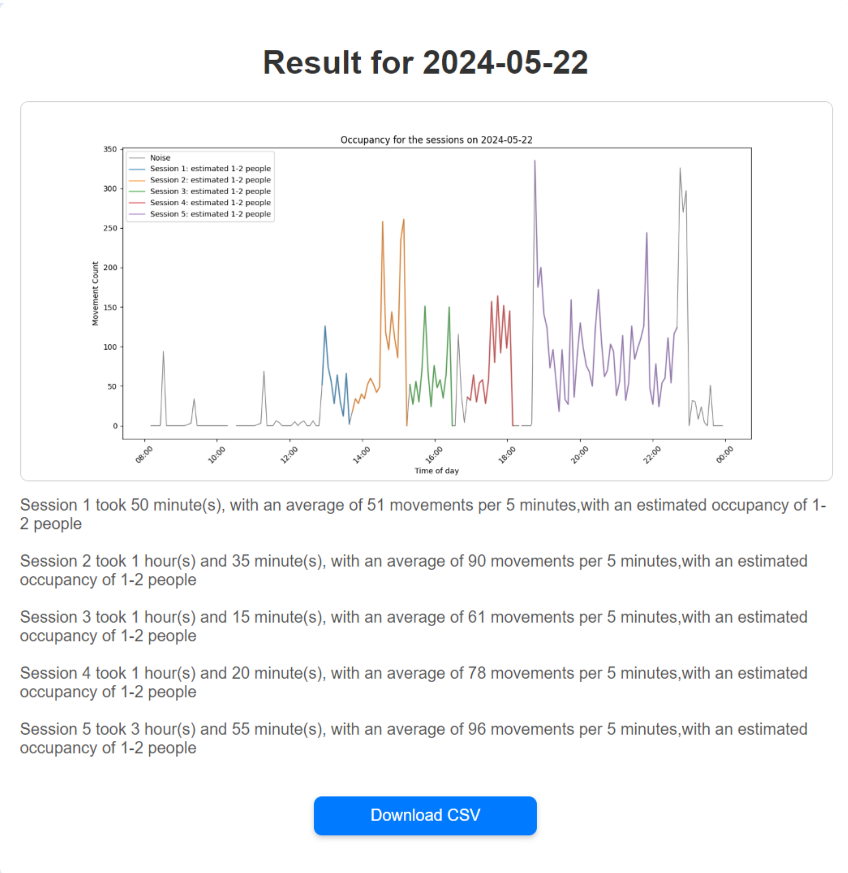

Figure 8: Example of an analysed day

4.2 Jupyter-notebook UI

The Jupyter-notebook provides a more simple interface where the user is asked to provide a file path, a date to analyse and the desired estimation mode 9. The user will receive the exact same result as the single day option in the web-based UI. Besides that, the user can also check the dates within the file in order to make their choice.

Figure 9: Jupyter notebook interface

5 Discussion

The primary goal of this project was to create a model that could accurately estimate the occupancy of the study spaces at the UB, purely based on the movement counts detected by the sensors. Besides that, the UB wanted to gain an insight on the use of those study spaces. Data was gathered using two distinct motion sensors, the Locus sensor, which is placed underneath a desk, and the Alta sensor, which is placed on the ceiling of collaboration rooms. Since both sensors are very different, two different conclusions were made for the sensors which led to two different solutions. This final section will give these conclusions and discusses ideas that could be used for later versions of the final product and other research opportunities.

5.1 Conclusion

5.1.1 Locus Sensor

First of all, the Locus sensors that were used in this project were placed under desks intended for three people. However, after experiments were conducted it was evident that the detection range of these sensors were too small to be able to accurately capture three people sitting next to each other. For a single person the sensors did work at a satisfactory level, so given these facts, two different solutions were proposed to combat this issue. The first option is for the UB to simply install more sensors. This can be done by installing a locus sensor for every single meter of study space. However, this solution could be very costly and a lot of work, both in installing but also maintaining and managing. The second option is to replace the Locus sensor by the *Omni sensor*. This sensor is made by the same company, so it should be easy to integrate. This sensor is able to capture a 360 °range of motion and it also has a larger detection range than the Locus sensor.

5.1.2 Alta Sensor

The Alta sensor, on the other hand, was able to provide useful data. This is supported by the fact a useful model was created which led to the final product. This model is able to estimate the occupancy of the collaboration rooms solely based on the motion counts the Alta sensor provides. This model estimates the occupancy on a study session bases as opposed to estimating the occupancy for every single data point, as the former is a more intuitive way of representing the occupancy of a study space. The model splits the data into sessions, removes outliers, estimates occupancy and provides a useful visualisation and summary of the sessions. Besides showing useful information an option to download the important information of the days in CSV-format was added, per request of the product owners. This was done to easily conduct more research with the acquired and processed data. This model is translated into an easy to use web-based user interface and a Jupyter-notebook. However, the final application does not work in real-time right now. Besides that, as of now an accuracy score can't be calculated because of the lack in target data. However, the model can be visually assessed by the test findings.

5.2 Further Improvements

While the final product provides a solid overview and estimation on the occupancy, the model behind it has room for further optimization. One of these optimizations is the occupancy estimation model. A crucial factor to improve this model is to gather more accurate target data that the model can be trained on. A realistic way to generate this target data is by installing low resolution, high frame-rate cameras that can accurately identify whether someone enters the room, without being able to identify that person. These types of cameras do not infringe on privacy rules and are relatively cheap in use (Amin et al. 2008). When a large set of target data has been gathered, the model can be continuously improved to give even better estimations on the occupancy.
Furthermore, the application could be extended to also be useful for estimating the occupancy for single desks in study spaces. This could then be showcased on a map, so it is clear which desks were occupied or not. However, to realise this the proposed alternative sensors need to be

installed first. These sensors also need to be tested to ensure they function as intended and can capture movement more accurately.
Besides the data and model on which the application leans, the application could also be improved in usability and accessibility. The first of these improvements would be an option to select a custom time frame, so the user can see the data for other options than just a single day or an entire file. Furthermore, the application could utilise built-in data so users do not need to provide the data themselves.
Finally, when the application and model are satisfactory to all of the above points, it can be expanded to other UvA locations so the students, staff and researchers there can profit from the same perks as the UB.

References

Amin, Imran et al. (July 2008). "Automated people-counting by using low-resolution infrared and visual cameras". In: *Measurement* 41, pp. 589–599. DOI: 10.1016/j.measurement.2007.02.010.

Officebooking (2024). *Officebooking*. Accessed: 2024-06-30. URL: https://officebooking.com/nl/.

officebooking (2024a). *Alta*. Accessed: 2024-06-30. URL: https://officebooking.com/nl/products/esca-alta/.

— (2024b). *Locus*. Accessed: 2024-06-30. URL: https://officebooking.com/nl/products/esca-locus/.

— (2024c). *Omni sensor*. Accessed: 2024-06-30. URL: https://officebooking.com/en/esca-omni-occupancy-sensor/l.

University Library (2024). *About the Library*. Accessed: 2024-06-30. URL: https://uba.uva.nl/en/contact/about-the-library/about-the-library.

The final product can be found here: Final product UB repository.