

Addressing the Challenges of COVID-19 Social Distancing Through Passive Wi-Fi and Ubiquitous Analytics: A Real World Deployment

Miguel Ribeiro, Nuno Nunes, Marta Ferreira, João Nogueira, Johannes Schöning, Valentina Nisi

► To cite this version:

Miguel Ribeiro, Nuno Nunes, Marta Ferreira, João Nogueira, Johannes Schöning, et al.. Addressing the Challenges of COVID-19 Social Distancing Through Passive Wi-Fi and Ubiquitous Analytics: A Real World Deployment. 18th IFIP Conference on Human-Computer Interaction (INTERACT), Aug 2021, Bari, Italy. pp.3-24, 10.1007/978-3-030-85616-8_1. hal-04196882

HAL Id: hal-04196882 https://inria.hal.science/hal-04196882

Submitted on 5 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Addressing the Challenges of COVID-19 Social Distancing Through Passive Wi-Fi and Ubiquitous Analytics: A Real World Deployment

Miguel Ribeiro^{1,2}, Nuno Nunes^{1,2}, Marta Ferreira¹, João Nogueira¹, Johannes Schöning^{2,3,4}, and Valentina Nisi^{1,2}

¹ Instituto Superior Técnico, University of Lisbon, Portugal
² ITI/LARSyS, Portugal
³ University of Bremen, Germany
⁴ University of St. Gallen, Switzerland

Abstract. During the COVID-19 pandemic, social distancing measures were employed to contain its spread. This paper describes the deployment and testing of a passive Wi-Fi scanning system to help people keep track of crowded spaces, hence comply with social distancing measures. The system is based on passive Wi-Fi sensing to detect human presence in 93 locations around a medium-sized European Touristic Island. This data is then used in website plugins and a mobile application to inform citizens and tourists about the locations' crowdedness with real-time and historical data. To understand how people react to this type of information, we deployed online questionnaires in situ to collect user insights regarding the usefulness, safety, and privacy concerns. Results show that users considered the occupancy data reported by the system as positively related to their perception. Furthermore, the public display of this data made them feel safer while travelling and planning their commute.

Keywords: Passive Wi-Fi \cdot Survey \cdot Mobile Application \cdot COVID-19

1 Introduction

Driven by the impact of the COVID-19 pandemic, governments worldwide were forced to impose lockdowns, promote work from home, and enforce other measures for citizens to keep social distancing slow-down the spread of the virus and flatten the curve to avoid overwhelming the healthcare systems. In this scenario, we saw the appearance of several technological applications emerging, ranging from symptom tracking [23,27] and pandemic planning [44] to digital contact tracing [30,31,41,43] and also reaching the field of information visualization platforms [11,46]. Especially in high-density urban areas, maintaining social distancing is intricate. After the first wave of the pandemic from February to May 2020, many countries in the northern hemisphere saw a decrease in infection rates leading to a return of many economic activities. Driven by the need to foster economic growth, many countries incentivized unconfinement measures, from mandatory use of masks and sanitation to enforcing capacity limitations in public spaces.

This paper describes a technology-based measure implemented in a touristic European Island after the first wave of the pandemic. The system evolved from an existing community based passive Wi-Fi infrastructure deployed as part of a tourism flow sensing research project [29]. The infrastructure uses a network of Wi-Fi routers to passively detect the presence of wireless-enabled devices in their range [36]. Through machine learning, the infrastructure estimates the presence and flow of people both historically and in real-time. Unlike existing systems (e.g., Google popular times), the system is open to the local community enabling the deployment of additional services, including ubiquitous analytics [39] adapted to the local context. Here, we describe how we adapted the original infrastructure to help citizens and visitors, including local businesses and health and tourism authorities, leveraging open information about people's presence and flow. The infrastructure was deployed on 93 specific points of interest (POIs), emphasizing touristic places and popular public locations like shopping centers.

Sensing human presence has been approached in several manners, from manual people counting, surveys to wearable trackers [7]. In the context of the pandemic, monitoring mobility at scale and in real-time was an important requirement. Several methods were attempted, from video processing [15,38], crowdsourcing smartphone applications [33], cell tower and call record, and nonintrusive wireless monitoring (NIWM) [5,21,37,42]. While most techniques require user knowledge or intervention in the data acquisition process, NIWM does not. Either via cell signals, Bluetooth, or Wi-Fi devices, the number of people in specific locations can be sensed automatically and anonymously using NIWM. However, this data is hard to validate, as users often do not know they are being monitored or are not actively involved in the data collection process. To overcome this challenge, the system described here collects human presence data anonymously and presents the results back to citizens transparent and accessible. In this paper, we present our design approach and discuss how users understand and react to this real-time information. We contribute to a deeper understanding of such a system's impact in helping people maintain social distancing and the broader implications in terms of accuracy of the methods employed and the safety and privacy concerns that such information could entail.

1.1 Contributions and Research Questions

The work described in this paper was the response to the regional call to action to help the community return to a "new normal". The passive Wi-Fi infrastructure already in place in the island [29] was capable of estimating the presence and flow of people in public spaces. The existing system provided an opportunity to adapt the information for citizens and tourists, to help them make informed decisions when planning and moving through public locations. The system was then adapted from a generic passerby counter to a system to reinforce the social distancing measures of COVID-19 situation in the following ways: i) increase the existing network of passive Wi-Fi sensors to 93 POIs covering not only touristic POIs but also essential public places such as shopping centers and squares; ii) developing a mobile application that could show this information to citizens and locals, together with COVID-19 safety measures and tourist information about POIs and attractions; and finally, iii) deploying an in situ survey to present the occupancy data and probe passerby about the accuracy of the displayed data, as well as the sense of safety, privacy, and satisfaction of the location. Through this research, the authors intend to answer the following research questions:

RQ1: Can we validate passive Wi-Fi data as a representation of occupancy of a POI for COVID-19 occupancy monitoring?

RQ2: Can occupancy data automatically collected via passive Wi-Fi increase the sense of safety and crowd awareness in public locations?

RQ3: Can we use this information to help citizens make informed decisions about when to visit public locations based on their occupation?

2 Related Work

This research is primarily motivated by two areas of prior work, combining non intrusive people sensing and data presentation to citizens to collect feedback. The area of people sensing has had in the past the attention of researchers mainly due to the unnecessary intervention of the people being sensed. The technologies and techniques vary and aim to target the masses with automated mechanisms that require little to no human intervention. Examples of this are found in video surveillance analysis, where the video feed of existing surveillance cameras can be processed to enable head counts by using convolutional neural networks to overcome the challenges of low resolution images and poor lighting [15]. Recently a study [38] was done using thermal imagery to tag people for COVID-19 with signs like high body temperature by automatically detecting the faces and the temperature of the people passing in front of a thermal camera.

Other efforts were done using big data of cell phone records [40,4] to detect mobility traces, interaction patterns and taking advantage of functionalities such as handover and cell (re)selection, which are used to maintain seamless coverage for mobile end-user equipment between cells. The study by Alsaeedy and Chong [2] uses those parameters to identify the regions at risk of infections with high human mobility by using the existing cellular network. Also, Ahmed in [1] proposed a similar framework to control the spread of COVID-19 using the cellular system by using the mobile identifiers of infected people and calculating the distances of other phones nearby in those cell towers to deliver that information to the clinical centers for close contacts of suspects.

Wi-Fi has been used in the pre-pandemic to detect gatherings and passerby, with studies being done in football games [5], universities campuses [3] and hospitals [37]. The motivations of many of these studies are diverse, some look at energy waste on scanning methods [3], realistic facility management and planning [37], while others looked at crowding factors, flock detection and waiting times, speed and frequent paths [21] and even social information like popularity of events (in the case of [4] singers in concerts). In the pandemic context, Wi-Fi

has also been used to detect the contagion risk in public transport [10] by using multiple sensors that monitor the air quality, temperature humidity and use Wi-Fi probes to detect the people present thus creating an index for contagion risk. Another solution that employs Wi-Fi probe requests is presented in [42] with the goal to estimate indoor crowd densities based on dynamic fingerprints based on signal strength and proposing a crowd density estimation solution based on the positioning algorithm. The authors made experiments in a laboratory class and in three public social activities in a footbridge, railway and subway demonstrating how this can be used to estimate crowd density. Wi-fi was also proposed in a framework to monitor vital respiratory signs of patients by measuring the Wi-Fi frames reflected by the human body standing between a dedicated emitter and receiver and applying bandpass filtering to extract respiratory components of the patients [25].

2.1 Feedback via voluntary walk-up systems

While user surveys and online forms are viable ways for citizens to contribute with their opinions, the idea of gathering this type of information in situ without the need for a human interviewer opens opportunities for new studies. HCI brings new possibilities of presenting and gathering information from the users by providing interactive and attractive interfaces to grasp the passerby's attention often done via public displays. A recurrent problem with this method is that the users are often presented with automatically sensed information, that may not always represent the reality, and the user may feel the need to give an opinion or to even challenge the presented data. A classic example of this is the presentation of real-feel weather conditions or environmental parameters.

The area of surveying is essential to gather the citizens perspectives about different subjects of our society and products. Several approaches have been done to assess user satisfaction or inquiry through interactive technologies [6,19,20,28]. One of the studies done in this area presents Roam.IO [18], which proposes a hybrid data approach combining data from IoT sensing with qualitative data contributed by citizens. The authors developed a public installation consisting of a tangible robot-like interface placed in passageways, to entice and encourage the public to walk-up and answer questions. The users are questioned to suggest what the data might represent and enrich it with subjective observations with success in user engagement with this apparatus.

2.2 Citizen science

Several attempts have tried to tackle the problem of gathering data through citizens in ubiquitous ways [45,17]. This led to a range of interactive technologies, focused on human-centred data collection. An example is VoiceYourView [45] which uses speech recognition and natural language processing to provide feedback of how citizens perceive and feel about the public spaces around them. Physical tangible mechanisms have also been used for public opinion gathering such as polling systems to urban voting and visualizations [24]. Another tangible approach was EmoBall [12] and Mood Squeezer [14] which allows for playfulness in office buildings where a lightweight technology was designed to ask people to reflect on their mood by squeezing a colored ball from a box set. VoxBox [16] presented a tangible system for gathering opinions in situ at an event through playful and engaging interaction via a large panel of interaction turning the questionnaire itself into a toy that made people want to fidget with it. CommunitySourcing [17] was also used to investigate the potential of community sourcing by designing, implementing and evaluating a vending machine that allowed users to earn credits by performing tasks using a touchscreen attached to the machine and physical rewards were dispensed through traditional vending mechanics. Similarly, SmallTalk [13] explored how tangible devices used as a toy with spinners, dials and buttons can be used to gather feedback from children in the context of a theater performance via a tangible survey system to capture what children thought of the performance they had seen. Recently Diethei [8] explored socio-psychological processes and motivations to share personal data during a pandemic.

3 Data and Deployment

This system's main components rely on the presence and flow sensing infrastructure, providing user information via dashboards, web plugins, a mobile app, and finally collecting in situ feedback. The sensing infrastructure is based on a network of passive Wi-Fi routers described below. This infrastructure then powers the two other components by enabling data to be processed and presented to the end-users. The presence and flow data is presented to different stakeholders in different formats, from a global dashboard mostly targeting authorities to web plugins (that can be included in websites) targeting local businesses, thus informing them of people's presence at different times. Then, the data is consumed by an official Tourism Board app that presents the list of safe locations to visit and at what times. Finally, the surveys are generated as interactive webpages that can be deployed on tablets in situ or accessed by mobile phones via a QR code. The survey is used to collect feedback from the users of those locations about the accuracy of the data, safety of the location and overall satisfaction.

3.1 Passive Wi-Fi data gathering

The presence data presented in the visualizations is gathered from a passive Wi-Fi platform [29,36]. This infrastructure was originally designed to support a community-based Wi-Fi tracking system to understand mobility at scale. The original system provided several interactive dashboards [34,35,32] to help communities easily run systematic analysis of tourists' mobility patterns in the destinations, contributing in new ways in visualizing spatio-temporal mobility data.

This system is composed of Wi-Fi routers spread across all parishes of the region, as shown in Figure 1. The data is collected in real-time by Wi-Fi routers spread across multiple locations covering all regions' parishes. The data collected



Fig. 1. Map of Madeira Islands (Portugal) showing the locations of the passive Wi-Fi routers deployed.

and sent to a central server includes a location identifier, a timestamp and an anonymous identifier (based on the MAC address) of each Wi-Fi enabled device present in a radius of approximately 80 meters from the router (depending on the location and obstacles). The processed results are made available via an API that allows developing standalone implementations, thus providing real-time information about the number of devices detected in the last hours, location meta-data, and historical hourly data that serves as a comparison to the current counts.

The locations were chosen based upon common points of interest selected in coordination with the local tourism bureau and include governmental buildings close to the street, tourism offices, plazas, restaurants, cafes, and viewpoints where the routers were installed in the nearest facility. The routers placed in each location needed an electrical connection and internet access, thus posing challenges in terms of coverage and location. They had to be placed indoors close to outlets and on a suitable site provided by the existing building.

A challenge in adapting the infrastructure for COVID-19 security capacity was asking each site to provide the maximum safety number for each space. To overcome both challenges, the research team worked closely with local authorities to find and test a reasonable match between the router's range and the spaces' capacity. This process was tested initially in a private beach club, which provided very good conditions to evaluate the sensing infrastructure's accuracy. The club was only accessible via a controlled entrance and provided daily counts of admissions per hour and the maximum number of people inside the facilities on any given hourly period. This data was used as ground truth for several weeks to evaluate the accuracy of the machine learning methods used to estimate both the number of entrances and the total number of people inside the facility from the number of devices detected by the router. The ground truth facility had an overall capacity of 500 people and it was monitored with a single router located in the main entrance building. While this shows the simplicity and ease of deployment of the sensing infrastructure, not all the POIs could rely on the same level of accuracy since, in many places, there was no access control or no structured capacity defined. In these cases, the research team judged from inspection of maps and local screening based on signal strength measures. Finally, in some POIs, it was impossible to make the assessment, and hence the research team resorted to 60% of the maximum pre-pandemic capacity.

3.2 Dashboard and Web Plugins

Motivated by the rapid spread of COVID-19, and inspired by the social sensing paradigm, the contributions of this section lie in the implementation of an interactive platform based on the Wi-Fi tracking infrastructure, to visualize information related to this enriched dataset of spatio-temporal data. In particular, we here present the visualizations related to the COVID-19 disease diffusion and its effect on human mobility in the islands and describe the interactive data visualization.

We designed and implemented a web-based interactive data visualization built using Web technologies, focusing on the visualizations related to the COVID-19 period, showing the data in percentages of the total for each location, in a way to preserve the counts of some routers that are located in private infrastructures. These data representations have been designed to present the human presence considering the period when the pandemic of COVID-19 was experienced in Portugal. In particular, as it is possible to see in Figure 2, the presented data show the percentage of users in a specific area, in comparison with the last week before the day the state of emergency was declared in Portugal and the first case in Madeira Islands.

The dashboard presents an interactive visualization of human presence data during the COVID-19 outbreak, and to compare this data with the patterns of previous weeks. The visualized data is collected exploiting social sensing, and in particular, a community-based passive Wi-Fi tracking infrastructure, deployed in the Madeira Archipelago.



Fig. 2. (Left): Presents the information grouped by municipality; (Right) Zoomed view presents the ungrouped information focused on a specific area.

In addition to this, a web widget was also created being able to be embedded in websites, (which was also used for the in situ survey shown in Figure 3) so that each institution that has a router in its facilities can also inform its visitors about the current status and historical trend for each hour. This widget is currently used in the Private Beach Club website informing their website visitors of their affluence in real-time.

3.3 Mobile Application

The final touchpoint consisted of a mobile application designed for the Health and Tourism authorities responsible for implementing the Islands' safe tourism system. This system relied on the airport (and port) PCR COVID-19 testing. The visitors could either upload the results of a COVID-19 PCR test to the app and then pass the health check at the airports' arrivals. Alternatively, they could perform a free PCR test on arrival and receive the test results via email. During the 14-day monitoring period, the visitors were also asked to fill out a voluntary epidemiological survey.

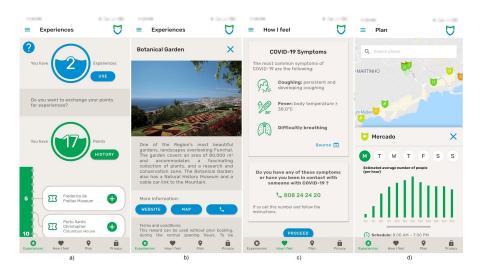


Fig. 3. Mobile application screens: a) **Experiences** - This view gives the user an overview of the current points collected and experiences exchange for points, as well as the history in a vertical chronological view; b) **How I Feel** - Shows the common user recommendations from the local health authorities and can call emergency numbers; c) **Plan** - This is where the user can explore the region in a map view and visualize the real-time occupancy of the locations marked by a shield that varies in color. Upon touching a site, the user is presented with historical occupancy for the different weekdays' data and shared information about the place, such as website, address, contacts, and schedule; d) **Privacy** - In this tab allows the user to view in a transparent manner how the information that is being presented is collected, and the ability to set notification permissions and geolocation (used to locate the user in the map).

Thus, the app was part of a more extensive safety system, which included testing all visitors, asking them to fill out an epidemiological survey and a daily follow up of COVID-19 related symptoms. To nudge visitors to perform the PCR test before arrival and as a way to incentivize daily epidemiological monitoring, the app also implemented a gamification mechanism that awarded users points each time they performed a COVID-19 safe activity.

Although the app's gamification component is beyond this paper's scope, it enabled a user to accumulate points by completing tasks and challenges of visiting safe locations. The points could then be exchanged for free experiences proposed by the local tourism office. The application interface starts by presenting the user with opening instructions for the first use, which helps the user navigate through the four main screens of the application and understand the points system, the challenges, and mechanics. Then the information is depicted in Figure 3 and categorized into four tabs.

Both the real-time map data and the historical occupation charts are powered by the passive Wi-Fi infrastructure that collects data anonymously in these locations.

3.4 In situ Survey

The in situ surveys were deployed in a total of four locations (described in more detail in Table 2). These surveys were used to present and collect information in selected POIs and survey the passerby's about the sense of safety and satisfaction of the location, and inquiry about privacy and safety concerns. This method also provided a way to crowdsource the accuracy of the information depicted, asking passersby to confirm the estimate and fine-tune the machine learning algorithms and the overall range and occupancy estimates.

The user interface was developed using web technologies suitable for responsive screen sizes and were deployed in a tablet to be answered via smartphone. The interface was designed as follows:

- To show the user the current location and current count (the system counts the number of devices in the Wi-Fi range and estimates the number of people based on a custom made algorithm) via a widget with an animated waveform gauge. The animation is supposed to draw attention to the screen (see Fig. 4, top left);
- Based on the real-time Wi-Fi counts, the interface shows an average hourly count history with the current hour highlighted (Fig. 4 Top);
- The interface shows the interactive survey. The survey gathers information about the users feeling about the occupancy levels of the location (as reported by the system interface), their safety perception, and the relevance of the real-time widget regarding their perceptions;
- The interface cycles through question panels when no interaction is recorded for 15 seconds. The cycling of the panels is designed to draw attention to the interactive screen;
- The system sends data to the database, when the user: answers all questions or does not interact for 15 seconds after having answered a question.

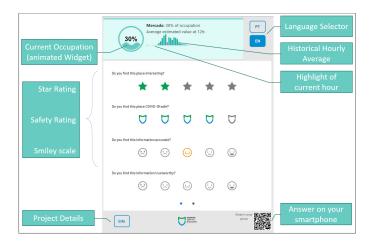


Fig. 4. In situ interactive survey Interface components diagram.

The survey had the option of two languages (English and Portuguese), has a QR-Code for the user to scan and answer the survey on their mobile phone thus avoiding physical touch with the tablet and has an "Info" button to show the project details and information about how the data is gathered and related privacy issues.

The survey questions are organised according to three kinds of answers, each one depicted by a specific icon: a star rating to rate the place interest, a shield to rate safety and a smiley for accuracy and trustworthiness score on a 5 point Likert scale We kept the number of questions low and succinct to maximize user engagement and a swift interaction with minimum effort. The questions are presented in table 1, grouped into two tabs of questions that cycled between them.

# Question	Rating Scale	Pannel	Connotation
Q1 Do you find this place interesting?	Star	1	Positive
Q2 Do you find this place COVID-19 safe?	Shield	1	Positive
Q3 Do you find this information accurate?	Smiley	1	Positive
Q4 Do you find this information trustworthy?	Smiley	1	Positive
Q5 Do you find this information useful?	Smiley	2	Positive
Q6 Do you find this information intrusive?	Smiley	2	Negative
Q7 Do you find this information improves safety?	Smiley	2	Positive
Q8 Do you find this information violates privacy?	Smiley	2	Negative

Table 1. In situ survey questions

11

3.5 Deployment

The in situ survey was deployed in four POIs and took place during the months of September to December of 2020. The locations were the research institute cafe (where it was piloted for a week), a private beach club, the gift shop of one of the most popular tourist attractions (viewpoint) and a Hotel lobby. In all of these places a tablet was used on a stand to show the relevant information and prompt the users with the questions.

The locations, described in table 2 were chosen for their characteristics of being locations that drew the attention of locals and tourists, and for their facilities that allowed an appreciate placement of the tablets indoors and in secure yet accessible spots in the premises.

Туре	Location	Limit	Nr Answers	Nr days
Research Institute	Restaurant/Bar	n/a	77	6
Touristic Attraction (Viewpoint)	Gift Shop	n/a	28	128
Beach Club	Entrance	500	90	131
Hotel	Reception	n/a	35	105
Total			230	_

Table 2. Location details

4 Results

In this section, we will present the results from the in situ survey, relative to how the users classified the information provided and perceived their safety at the locations. Besides, we present the results of modeling the Wi-Fi counts with ground truth data in one site to estimate the real number of people present.

4.1 Passive Wi-Fi ground truth validation

One of the four sites where the in situ survey was deployed, the private beach club, was used to provide ground truth to validate the data collected from the rest of the nodes of the passive Wi-Fi system and remodel the data to estimate the real number of people in the premises. To support the unconfinenment of people after the first wave of COVID -19, the regional authorities asked all public sites to limit occupancy. For this specific site, the number of people inside the premises was limited to 500. Since the access was controlled electronically, this site provided unique conditions to evaluate the accuracy of the Wi-Fi sensor. The ground truth data corresponds to 356 data points covering 44 days between June and August. The whole premise was covered by a single Wi-Fi sensor placed in the main building after the entrance.

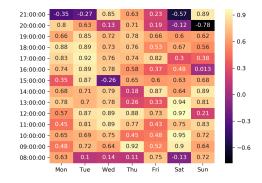


Fig. 5. Weekday and hourly correlations between the Wi-Fi counts and ground truth.

While the Wi-Fi counts help to show the peak times of occupation and the overall curve along the days of the week, those numbers are not exact counts. This is mainly due to people out of the Wi-Fi range having their wireless turned off, among other factors. To estimate the real number of people in the premises from the real-time Wi-Fi counts, we compared and created data models with the official hourly number of entrances and exits controlled via the electronic access control system. The time slot used for the comparisons was 60 minutes. For each hour and weekday, we computed the correlation between these variables. We tried to estimate the real number of people through data modeling. The data gathered in the private beach club shows strong Persons' correlations with the ground truth data. The correlation for the dataset is 0.64, and Figure 5 shows the breakdown in weekdays and hours. In this analysis, each correlation is done using 6 data points. The low values in some isolated hours correspond to the beginning and end of the day, where abnormal movements may have occurred during the procedures of opening and closing of the facilities that did not provide correlations as good as the remaining hours.

We also attempted to model the number of actual people entering and exiting the premises against the router counts with several regression methods. The features were extracted by grouping the data hourly for each weekday. Each

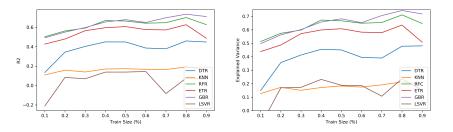


Fig. 6. Evolution of R2 (left) and explained variance (right) with incremental steps of train/test percentages for seven methods.

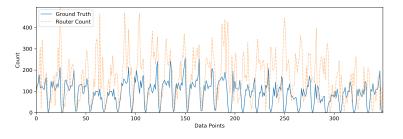


Fig. 7. Beach club Ground truth counts vs router counts.

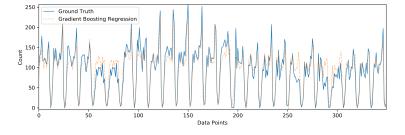


Fig. 8. Modeled router counts of the beach club to fit the ground truth data.

row represented one hour of data on a specific weekday from the location. The number of hours of ground truth varied, as also did the closing times. The total number of days was 44, ranging from 2020-06-15 to 2020-07-29. This resulted in 590 data points with the following features: *hour, weekday, count.* The Scikit-learn python library was used to test the following common methods used in regression:

- Decision Tree Regressor (DTR)
- K-Neighbors Regressor (KNR) k=100
- Random Forest Regressor (RFR) Estimators=100
- Extra Trees Regressor (ETR) Estimators=100
- Gradient Boosting Regressor (BGR) Learning Rate=0.1, Estimators=100
- Linear Support Vector Regressor (LSVR) = Tolerance=1e-4
- Epsilon Support Vector Regressor (SVR) Degree=3, Epsilon=0.1

The parameter sweep was done for the train/test ratio, ranging from 10% to 90%, with step increments of 10%. These classifiers were run 1000 times for each of the steps and the average accuracy for each classifier was taken for the results shown in Figure 6. The method that shows higher results when using the same dataset for trains and test (train size=1) is the Gradient Boost Regression. Using this method, we plotted in Figure 7 the original data points of router counts and ground truth, and in Figure 8 the regression made using the Random Forest Regressor, with values of R2 and explained variance of 0.92.

These results show that the routers data can be reliably used to predict the real number of passersby in certain locations by modeling for each location, or

Table 3. Statistical analysis of replies (Q1-Interesting; Q2-Safe; Q3-Accurate; Q4-Trustworthy; Q5-Useful; Q6-Intrusive; Q7-Improves safety; Q8-Violates Privacy)

Parameter	Q1	Q2	Q3	$\mathbf{Q4}$	Q5	Q6	Q7	Q8
Ν	185	182	174	163	172	155	155	149
Average	4.58	4.41	4.17	4.26	4.40	3.46	4.43	3.29
Std. Dev.	0.95	1.09	1.19	1.13	1.08	1.60	1.01	1.69
Median	5	5	5	5	5	4	5	4

location typology. The accuracy depends on how the router is installed, its range, how many people leave Wi-Fi on, as well as being influenced by the random MAC addresses, however, as we can see, it is still possible to create accurate models that predict well the occupancy of a location.

4.2 In situ survey

During the survey deployment, we received a total of 230 responses. All of the questions were formulated affirmatively except two (Q6 and Q8), concerning privacy and intrusiveness. The user's answers can be interpreted as the higher the scales' rating, the better.

The statistics for the 230 replies to the eight questions are shown in table 3. It shows that Q6(privacy) and Q8(intrusiveness) have more spread out standard deviations, while the remaining questions scored high averages, showing that the public information derived from the passive Wi-Fi data was correlated with the location's perceived occupancy. In general the users reported interest in the site, and felt COVID-19 safe. Since the survey could be abandoned without answering all of the questions, the total number of users that replied to all questions was 130. The answers regarding interest in the place (Q1), perceived safety in the place, accuracy of info, trustworthiness, usefulness, and improved safety of the

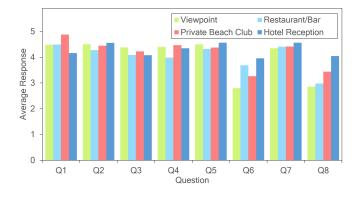


Fig. 9. Average question results across the different locations.

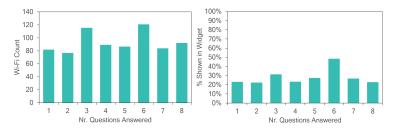


Fig. 10. Number of questions answered by Wi-Fi count(left) and percentage shown in the widget(right).

place through the info provided (Q2,Q3,Q4,Q5,Q7), are all quite positive (between 4 and 5 Likert scale rating). The questions regarding intrusiveness of the information (Q6) and violation of privacy (Q8) recorded lower scores, compared to the previous answers, but still quite high, (between 3 and 4), according to the Likert scale. Since the questions were formulated in a negative form, high scores in these two questions indicate users perceptions of the information being intrusive and privacy violation.

Figure 9 shows the score distribution of the questions across the four locations is consistent across all sites, except for questions Q6 and Q8 regarding privacy and intrusiveness. For these questions, the Hotel and Bar show users being more concerned about privacy (Q8) and intrusiveness (Q6). To note that in all questions, the maximum standard deviation (SD) was 1.69 and only 2 questions had SDs above 1.19 in a 5 point Likert scale, which implies that most answers in these 2 questions ranged between ~2.4 points of the scale, thus not being able to extract conclusions as strong and consistent as the remaining questions.

In Figure 10 we present the number of questions answered compared to the number of Wi-Fi devices detected in the past 60 minutes to the reply and the occupation percentage shown in the widget. The results reveal no direct increase or decrease resulting from the percentage shown and the number of questions answered. Six, was the most prevalent number of questions answered when the location showed more occupancy.

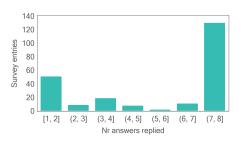


Fig. 11. Distribution of number of questions answered by each user.

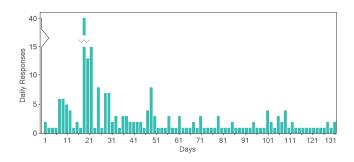


Fig. 12. Number of daily responses throughout the deployment (Broken Y axis).

Figure 11 shows the distribution of the number of questions answered by each user, revealing that the vast majority of users replied to all eight questions or just one single question before abandoning the survey. The deployment spanned from September to December of 2020. The number of daily responses varied considerably, as shown in Figure 12, with a higher frequency happening in the initial days and then rapidly fading out through the deployment. To note that across the four locations, at least one response was recorded every day. We also plotted the replies for questions 1-5 and 7 against the percentage shown in the widget in Figure 13. These were the questions with positive connotation aimed to understand the user's satisfaction with the location and the perception of accuracy of the data being presented.

Lastly, the language in which the questionnaires were answered may differentiate the users that answered from local and foreigners. The results presented in Fig. 14 (left) shows that the majority of survey entries were done in portuguese, and that division is maintained across the users of the other locations, with the exception of the hotel reception, in which the majority of entries were done in english as seen in Fig. 14 (right).

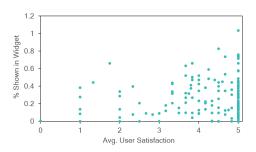


Fig. 13. Average reply for questions 1-5 and 7 (positive connotation) against the percentage shown in the widget.

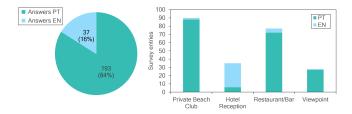


Fig. 14. Number of survey entries per language (left) and number of questions answered per survey language (right).

5 Discussion

As a low-cost technological solution to help unconfinement measures during the pandemic, the infrastructure described here provided some useful results and discussion points. Monitoring and modeling people's behavior is always a sensitive issue. The passive Wi-Fi approach is both anonymous and practical for providing real-time information about the people's presence, occupancy and flow. Previous research showed this method was effective in detecting people's presence and flow while preserving privacy [29,36]. Although Wi-Fi devices expose IDs that could potentially be used to track the owners, our dataset clearly shows that this is impractical with the high number (96%) of randomized IDs exposed in recent iOS and Android implementations. Although, to increase privacy, the system does not store device IDs, our past data shows the randomization percentage has risen significantly in the last years (56% since 2016). This is an important privacy by design feature, making passive Wi-Fi more trustworthy than other alternatives, e.g., using cameras [47] or access control systems or even active Wi-Fi logging [26]. Although systems like Google popular times provide similar information, these datasets are not publicly available and therefore, are not a practical solution for communities such as those described here. As drawbacks, the passive Wi-Fi system requires an infrastructure to operate (power, internet connection, backend for ML algorithms and provide real-time and historical data). While our results show a strong correlation between the passive Wi-Fi data and the ground truth, the system is naturally prone to inaccuracies. The location (including obstacles), the limited range, and dependency on having Wi-Fi enabled can naturally influence the system's accuracy. One could argue that such potential inaccuracies are a built-in feature as they end up increasing privacy protection (e.g., acting as a non-intentional mechanism to implement differential privacy [9]). Despite the limitations, our results clearly show that passive Wi-Fi is a viable alternative for capacity monitoring in the wild, answering RQ1.

Besides, the in situ surveys were a mechanism to understand people's reaction of this information. First, the survey results confirmed that users perceived the occupancy data displayed to be quite accurate, confirming from a user perspective our RQ1 (Q3-Q5). The results show that users were overall pleased with the information provided. From the answers regarding Q7 (safety improval) the average result of 4.43 shows that the users perceived that the system presented

improved the citizens feeling of safety in that location while also finding the place COVID-safe (Q2), which also comes from a combination of the measures taken by the local authorities.

Secondly, our deployment confirms that in situ surveying is prone to the novelty effect, exciting users in the beginning, while not supporting long term engagement of the same number of users [22]. However, it did continue to engage some users over the whole deployment (4 months) since we recorded at least one answer per location over the period of 127 days. Finally, concerning privacy and intrusiveness (Q6 and Q8), the users express some negative concerns about the system as the Likert's scores are quite high (3 to 4) and the question phrased negatively. Further probing of the users would be necessary to understand their concerns in more details.

Future work should address differences between in situ and mobile app use of this same information. The mobile application has been released, being promoted mainly via touristic channels and was installed by over 9000 users in Google and Apple stores (Jan 2021 to Apr 2021). From the usage data, and the time spend in each screen, the users explored on average 13 locations of the interactive map and opened the hourly occupation charts, thus showing that the application was used to inform users of the real-time occupation in several locations. In addition, given the positive feedback of the data accuracy (Q3) and trustworthiness (Q4) from the in situ survey (the same rates shown in the app), the data shown in the mobile application enabled the users to make informed decisions about their commute plans around the region as per RQ3.

6 Conclusion

With this work, describe and evaluate the deployment of three components that aim to inform citizens about the real-time occupation of public spaces by using a passive Wi-Fi system. The challenge of estimating the real number of people from the Wi-Fi counts was modeled for a use case with accurate ground truth, that provides a solid base for modeling data accurately. The data modeling from the use case used for modeling here set the ground to update the rest of the Wi-Fi locations to be used in that way, opening the possibility of creating models for locations with similar characteristics once the basic info is provided, such as the capacity of the places, and the accuracy of how much of the passerby does the system capture, while adjusting the routers so to cover all ground necessary to estimate the number of people in that location.

The positive feedback of the survey system encourages future work to turn the web dashboard presented into a larger public display to be placed in the cities and inform citizens in real time about the presence of people in locations nearby with the updated models for each location. While the mobile application will also be updated to include quick user surveys that prompt users with questions about the experience of using the application and the general feel of the visited locations as well as feedback about the presented data.

References

- Ahmed, A.M.: Designing a framework to control the spread of covid-19 by utilizing cellular system. Kurdistan Journal of Applied Research pp. 146–153 (2020). https://doi.org/10.24017/covid.16
- Alsaeedy, A.A.R., Chong, E.K.P.: Detecting regions at risk for spreading COVID-19 using existing cellular wireless network functionalities. IEEE Open Journal of Engineering in Medicine and Biology 1, 187–189 (2020). https://doi.org/10.1109/OJEMB.2020.3002447, conference Name: IEEE Open Journal of Engineering in Medicine and Biology
- Baniukevic, A., Jensen, C., Lu, H.: Hybrid indoor positioning with wi-fi and bluetooth: Architecture and performance. In: 2013 IEEE 14th International Conference on Mobile Data Management (MDM). vol. 1, pp. 207–216 (2013). https://doi.org/10.1109/MDM.2013.30
- 4. Böhmer, M., Hecht, B., Schöning, J., Krüger, A., Bauer, G.: Falling asleep with angry birds, facebook and kindle: a large scale study on mobile application usage. In: Proceedings of the 13th international conference on Human computer interaction with mobile devices and services. pp. 47–56 (2011)
- Bonné, B., Barzan, A., Quax, P., Lamotte, W.: WiFiPi: Involuntary tracking of visitors at mass events. In: 2013 IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM). pp. 1–6 (2013). https://doi.org/10.1109/WoWMoM.2013.6583443
- Brignull, H., Rogers, Y.: Enticing people to interact with large public displays in public spaces. In: Interact. vol. 3, pp. 17–24 (2003)
- Buettner, M., Prasad, R., Philipose, M., Wetherall, D.: Recognizing daily activities with rfid-based sensors. In: Proceedings of the 11th international conference on Ubiquitous computing. pp. 51–60 (2009). https://doi.org/10.1145/1620545.1620553
- Diethei, D., Niess, J., Stellmacher, C., Stefanidi, E., Schöning, J.: Sharing heartbeats: Motivations of citizen scientists in times of crises. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3411764.3445665, bluehttps://doi.org/10.1145/ 3411764.3445665
- Dwork, C.: Differential privacy: A survey of results. In: International conference on theory and applications of models of computation. pp. 1–19. Springer (2008). https://doi.org/10.1007/978-3-540-79228-41
- Fatih, Ş.İ., Gökhan, A.N., Panić, S., Stefanović, Č., Yağanoğlu, M., Prilinčević, B.: Covid-19 risk assessment in public transport using ambient sensor data and wireless communications. Bulletin of Natural Sciences Research 10(2), 43–50 (2020). https://doi.org/10.5937/bnsr10-29239
- Florez, H., Singh, S.: Online dashboard and data analysis approach for assessing COVID-19 case and death data. F1000Research 9, 570 (2020). https://doi.org/10.12688/f1000research.24164.1
- Fuentes, C., Rodríguez, I., Herskovic, V.: Emoball: a study on a tangible interface to self-report emotional information considering digital competences. In: Ambient Intelligence for Health. pp. 189–200. Springer (2015). https://doi.org/10.1007/978-3-319-26508-7₁9
- 13. Gallacher, S., Golsteijn, C., Rogers, Y., Capra, L., Eustace, S.: Smalltalk: using tangible interactions to gather feedback from children. In: Proceedings of the TEI'16:

Tenth International Conference on Tangible, Embedded, and Embodied Interaction. pp. 253–261 (2016). https://doi.org/10.1145/2839462.2839481

- Gallacher, S., O'Connor, J., Bird, J., Rogers, Y., Capra, L., Harrison, D., Marshall, P.: Mood squeezer: lightening up the workplace through playful and lightweight interactions. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. pp. 891–902 (2015). https://doi.org/10.1145/2675133.2675170
- Gao, C., Li, P., Zhang, Y., Liu, J., Wang, L.: People counting based on head detection combining adaboost and CNN in crowded surveillance environment. Neurocomputing 208, 108–116 (2016). https://doi.org/10.1016/j.neucom.2016.01.097
- Golsteijn, C., Gallacher, S., Koeman, L., Wall, L., Andberg, S., Rogers, Y., Capra, L.: Voxbox: A tangible machine that gathers opinions from the public at events. In: Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction. pp. 201–208 (2015). https://doi.org/10.1145/2677199.2680588
- Heimerl, K., Gawalt, B., Chen, K., Parikh, T., Hartmann, B.: Communitysourcing: engaging local crowds to perform expert work via physical kiosks. In: Proceedings of the SIGCHI conference on human factors in computing systems. pp. 1539–1548 (2012). https://doi.org/10.1145/2207676.2208619
- Houben, S., Bengler, B., Gavrilov, D., Gallacher, S., Nisi, V., Nunes, N.J., Capra, L., Rogers, Y.: Roam-IO: Engaging with people tracking data through an interactive physical data installation. In: Proceedings of the 2019 on Designing Interactive Systems Conference. pp. 1157–1169. DIS '19, Association for Computing Machinery (2019). https://doi.org/10.1145/3322276.3322303
- Houben, S., Weichel, C.: Overcoming interaction blindness through curiosity objects. In: CHI '13 Extended Abstracts on Human Factors in Computing Systems. pp. 1539–1544. CHI EA '13, Association for Computing Machinery (2013). https://doi.org/10.1145/2468356.2468631
- Ju, W., Sirkin, D.: Animate objects: How physical motion encourages public interaction. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) Persuasive Technology. pp. 40–51. Lecture Notes in Computer Science, Springer (2010). https://doi.org/10.1007/978-3-642-13226-1₆
- Kjærgaard, M.B., Wirz, M., Roggen, D., Tröster, G.: Mobile sensing of pedestrian flocks in indoor environments using WiFi signals. In: 2012 IEEE International Conference on Pervasive Computing and Communications. pp. 95–102 (2012). https://doi.org/10.1109/PerCom.2012.6199854
- 22. Koch, M., von Luck, K., Schwarzer, J., Draheim, S.: The novelty effect in large display deployments-experiences and lessons-learned for evaluating prototypes. In: Proceedings of 16th European Conference on Computer-Supported Cooperative Work-Exploratory Papers. European Society for Socially Embedded Technologies (EUSSET) (2018). https://doi.org/10.18420/ecscw2018₃
- Koehlmoos, T.P., Janvrin, M.L., Korona-Bailey, J., Madsen, C., Sturdivant, R.: Covid-19 self-reported symptom tracking programs in the united states: Framework synthesis. Journal of medical Internet research 22(10), e23297 (2020). https://doi.org/10.2196/23297
- Koeman, L., Kalnikaité, V., Rogers, Y.: "Everyone is talking about it!": A distributed approach to urban voting technology and visualisations. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems. pp. 3127–3136 (2015). https://doi.org/10.1145/2702123.2702263

Addressing Challenges of COVID-19 Social Distancing with Passive Wi-Fi

21

- Li, F., Valero, M., Shahriar, H., Khan, R.A., Ahamed, S.I.: Wi-COVID: A COVID-19 symptom detection and patient monitoring framework using WiFi. Smart Health 19, 100147 (2021). https://doi.org/10.1016/j.smhl.2020.100147
- Meneses, F., Moreira, A.: Large scale movement analysis from WiFi based location data. In: 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN). pp. 1–9 (2012). https://doi.org/10.1109/IPIN.2012.6418885
- 27. Menni, C., Valdes, A.M., Freidin, M.B., Sudre, C.H., Nguyen, L.H., Drew, D.A., Ganesh, S., Varsavsky, T., Cardoso, M.J., El-Sayed Moustafa, J.S., Visconti, A., Hysi, P., Bowyer, R.C.E., Mangino, M., Falchi, M., Wolf, J., Ourselin, S., Chan, A.T., Steves, C.J., Spector, T.D.: Real-time tracking of self-reported symptoms to predict potential COVID-19. Nature Medicine **26**(7), 1037–1040 (2020). https://doi.org/10.1038/s41591-020-0916-2, number: 7 Publisher: Nature Publishing Group
- Müller, J., Alt, F., Michelis, D., Schmidt, A.: Requirements and design space for interactive public displays. In: Proceedings of the 18th ACM international conference on Multimedia. pp. 1285–1294 (2010). https://doi.org/10.1145/1873951.1874203
- Nunes, N., Ribeiro, M., Prandi, C., Nisi, V.: Beanstalk: a community based passive wi-fi tracking system for analysing tourism dynamics. In: Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems. pp. 93–98 (2017). https://doi.org/10.1145/3102113.3102142
- 30. Oswald, M., Grace, J.: The covid-19 contact tracing app in england and 'experimental proportionality'. SSRN Electronic Journal (2020). https://doi.org/10.2139/ssrn.3632870
- Park, Y.J., Choe, Y.J., Park, O., Park, S.Y., Kim, Y.M., Kim, J., Kweon, S., Woo, Y., Gwack, J., Kim, S.S., Lee, J., Hyun, J., Ryu, B., Jang, Y.S., Kim, H., Shin, S.H., Yi, S., Lee, S., Kim, H.K., Lee, H., Jin, Y., Park, E., Choi, S.W., Kim, M., Song, J., Choi, S.W., Kim, D., Jeon, B.H., Yoo, H., Jeong, E.K., COVID-19 National Emergency Response Center, Epidemiology and Case Management Team: Contact tracing during coronavirus disease outbreak, south korea, 2020. Emerging Infectious Diseases 26(10), 2465–2468 (2020). https://doi.org/10.3201/eid2610.201315
- 32. Prandi, C., Nisi, V., Ribeiro, M., Nunes, N.: Sensing and making sense of tourism flows and urban data to foster sustainability awareness: a real-world experience. Journal of big Data 8(1), 1–25 (2021). https://doi.org/10.1186/s40537-021-00442w
- Ram, N., Gray, D.: Mass surveillance in the age of COVID-19. Journal of Law and the Biosciences 7 (2020). https://doi.org/10.1093/jlb/lsaa023
- 34. Redin, D., Vilela, D., Nunes, N., Ribeiro, M., Prandi, C.: Vitflow: a platform to visualize tourists flows in a rich interactive map-based interface. In: 2017 Sustainable Internet and ICT for Sustainability (SustainIT). pp. 1–2. IEEE (2017). https://doi.org/10.23919/SustainIT.2017.8379814
- Ribeiro, M., Nisi, V., Prandi, C., Nunes, N.: A data visualization interactive exploration of human mobility data during the covid-19 outbreak: a case study. In: 2020 IEEE Symposium on Computers and Communications (ISCC). pp. 1–6. IEEE (2020). https://doi.org/10.1109/ISCC50000.2020.9219552
- Ribeiro, M., Nunes, N., Nisi, V., Schöning, J.: Passive wi-fi monitoring in the wild: a long-term study across multiple location typologies. Personal and ubiquitous computing pp. 1–15 (2020). https://doi.org/10.1007/s00779-020-01441-z
- 37. Ruiz-Ruiz, A.J., Blunck, H., Prentow, T.S., Stisen, A., Kjærgaard, M.B.: Analysis methods for extracting knowledge from large-scale WiFi monitoring to inform building facility planning. In: 2014 IEEE International Conference

on Pervasive Computing and Communications (PerCom). pp. 130–138 (2014). https://doi.org/10.1109/PerCom.2014.6813953

- Said, M., Samuel, M., Shannan, N., Faizah, M., Bashir, Dodo, Y.: Novel visionbased thermal people counting tool for tracking infected people with viruses like COVID-19. Journal of Adv Research in Dynamical Control Systems 12, 1115–1119 (2020). https://doi.org/10.5373/JARDCS/V12SP7/20202210
- 39. Shaw, P., Mikusz, M., Nurmi, P., Davies, N.: Tacita: A privacy preserving public display personalisation service. In: Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. pp. 448–451 (2018). https://doi.org/10.1145/3267305.3267627
- 40. Sohn, T., Varshavsky, A., LaMarca, A., Chen, M.Y., Choudhury, T., Smith, I., Consolvo, S., Hightower, J., Griswold, W.G., de Lara, E.: Mobility detection using everyday GSM traces. In: Dourish, P., Friday, A. (eds.) UbiComp 2006: Ubiquitous Computing. pp. 212–224. Lecture Notes in Computer Science, Springer Berlin Heidelberg (2006)
- 41. Stevens, H., Haines, M.B.: TraceTogether: Pandemic response, democracy, and technology. East Asian Science, Technology and Society 14(3), 523–532 (2020). https://doi.org/10.1215/18752160-8698301
- 42. Tang, X., Xiao, B., Li, K.: Indoor crowd density estimation through mobile smartphone wi-fi probes. IEEE Transactions on Systems, Man, and Cybernetics: Systems 50(7), 2638–2649 (2020). https://doi.org/10.1109/TSMC.2018.2824903, conference Name: IEEE Transactions on Systems, Man, and Cybernetics: Systems
- 43. Vedaei, S.S., Fotovvat, A., Mohebbian, M.R., Rahman, G.M.E., Wahid, K.A., Babyn, P., Marateb, H.R., Mansourian, M., Sami, R.: COVID-SAFE: An IoT-based system for automated health monitoring and surveillance in post-pandemic life. IEEE Access 8, 188538–188551 (2020). https://doi.org/10.1109/ACCESS.2020.3030194, conference Name: IEEE Access
- 44. Whitelaw, S., Mamas, M.A., Topol, E., Spall, H.G.C.V.: Applications of digital technology in COVID-19 pandemic planning and response. The Lancet Digital Health 2(8), e435 – e440 (2020). https://doi.org/https://doi.org/10.1016/S2589-7500(20)30142-4
- 45. Whittle, J., Simm, W., Ferrario, M.A., Frankova, K., Garton, L., Wood-cock, A., Nasa, B., Binner, J., Ariyatum, A.: Voiceyourview: collecting real-time feedback on the design of public spaces. In: Proceedings of the 12th ACM international conference on Ubiquitous computing. pp. 41–50 (2010). https://doi.org/10.1145/1864349.1864358
- 46. Wissel, B.D., Van Camp, P.J., Kouril, M., Weis, C., Glauser, T.A., White, P.S., Kohane, I.S., Dexheimer, J.W.: An interactive online dashboard for tracking COVID-19 in u.s. counties, cities, and states in real time. Journal of the American Medical Informatics Association 27(7), 1121–1125 (2020). https://doi.org/10.1093/jamia/ocaa071
- 47. Zhao, X., Delleandrea, E., Chen, L.: A people counting system based on face detection and tracking in a video. In: 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance. pp. 67–72. IEEE (2009). https://doi.org/10.1109/AVSS.2009.45