

---

# Lessons Learned Using Wi-Fi and Bluetooth as Means to Monitor Public Service Usage

**Lu Bai**

Department of Computer  
Science  
University of Sheffield, UK  
lu.bai@sheffield.ac.uk

**Neil Ireson**

Department of Computer  
Science  
University of Sheffield, UK  
n.ireson@sheffield.ac.uk

**Suvodeep Mazumdar**

Department of Computer  
Science  
University of Sheffield, UK  
s.mazumdar@sheffield.ac.uk

**Fabio Ciravegna**

Department of Computer  
Science  
University of Sheffield, UK  
f.ciravegna@shef.ac.uk

**Abstract**

Facets of urban public transport such as occupancy, waiting times, route preferences are essential to help deliver improved services as well as better information for passengers to plan their daily travel. The ability to automatically estimate passenger occupancy in near real-time throughout cities will be a step change in the way public service usage is currently estimated and provide significant insights to decision makers. The ever-increasing popularity and abundance of mobile devices with always-on Wi-Fi/Bluetooth interfaces makes Wi-Fi/Bluetooth sensing a promising approach for estimating passenger load. In this paper, we present a Wi-Fi/Bluetooth sensing system to detect mobile devices for estimating passenger counts using public transport. We present our findings on an initial set of experiments on a series of bus/tram journeys encapsulating different scenarios over five days in a UK metropolitan area. Our initial experiments show promising results and we present our plans for future large-scale experiments.

**Author Keywords**

Wi-Fi/Bluetooth sensing; low cost; passenger load estimation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).  
UbiComp/ISWC'17 Adjunct, September 11–15, 2017, Maui, HI, USA  
© 2017 Copyright is held by the owner/author(s). Publication rights licensed to ACM.  
ACM 978-1-4503-5190-4/17/09...\$15.00  
<https://doi.org/10.1145/3123024.3124417>

## **ACM Classification Keywords**

C.3 [SPECIAL-PURPOSE AND APPLICATION-BASED SYSTEMS].

## **Introduction**

As cities grow, a paradox results: the demand for personal mobility and the mobility of goods and services increases, but the possibilities of meeting those demands diminish, especially through public services. This occurs particularly where spatial patterns of growth have resulted in urban sprawl and the lowering of average densities, increasing average journey lengths and congestion on overburdened infrastructure [1] [2]. The per-unit costs of transport services (including the provision of roads) increases, as the level of service declines. In a period of dwindling public funds and drastic reduction in public service availability, however, the provision of public services is likely to reduce rather than expand. There is therefore an urgent need to rationalise public services and infrastructure so to serve communities in an efficient and effective manner at a cost that is affordable and sustainable, while protecting vulnerable and disadvantaged communities. A major bottleneck in the provision of targeted public services (e.g. buses, cycle routes, etc.) and active traveling is the precise and timely quantification and understanding of communities' needs or the potential take up of new initiatives.

Traditionally, public transport routes and decisions have been made based on manual collection of passenger load information [3]. However, conducting such surveys is a highly expensive, labour intensive and time consuming process and hence, only conducted in limited numbers. This leads to weaknesses in the

quantitative evidence base that is required to objectively design targeted services. While bus service quality is evaluated by several factors such as frequency, waiting times, cost, cleanliness, travel time etc. [4] [5], vehicle occupancy is the most commonly used [6].

In order to respond to the need for accurate real time bus occupancy information, there is a need for passive sensing approaches. Automatic passenger counting is gradually becoming one of the more popular solutions with the emergence of surveillance camera/image based monitoring techniques [7] [8]. Such techniques rely on the camera's field of vision and hence are prone to be inaccurate with the presence of obstruction, poor light or overcrowded in the vehicle [9]. Other sensing technologies have also been explored so far e.g. using doorway infrared sensor [10], RFID-sensor [11], smart ticket [12] etc.

While camera/image based monitoring techniques and IR sensors provide means for reducing expensive manual surveys, they still suffer from an inability to accurately identify distinct individuals, and rely on coarse-grained spatial information. Smartphone based Bluetooth or Wi-Fi sensing is a promising alternative that can help alleviate these challenges [13] [14]. With an increasing adoption of Bluetooth Low Energy (LE), sensing can be further refined with improved higher coverage.

We discuss our approach toward developing a reliable, non-intrusive and passive mechanism to estimate passenger load on public transport (bus) and understanding waiting time patterns at bus stops. In this paper, we present a system that employs a passive



Figure 1: Raspberry Pi based static sensor System Hardware

sensing approach for monitoring Bluetooth and Wi-Fi probe requests to estimate the number of people within the sensor's vicinity. One major feature of our approach is the potential for application over large scale. In order to gather a holistic understanding of public transport service within wide regions and cities, hundreds of sensors are needed to be deployed in public spaces, stations and buses. Our solution is low cost, simple to deploy, lightweight and portable; experiments show it able to allow reliable user quantification.

The paper is structured as follows: we initially present related work on the field; we then present the system we have developed. We discuss a set of experiments conducted as a part of a preliminary feasibility study to understand various aspects of our approach. We then present our results and conclude the paper with some discussions on future work.

### Related Work

Tracking Bluetooth devices for monitoring citizens and public has recently seen an emerging interest [15] [16]. In comparison to Bluetooth, Wi-Fi passive sensing provides a greater coverage of the users, and has also been explored in the past. For example, [17] used a Wi-Fi based sensing solution to detect and track users. The authors reported achieving an accuracy of more than 75% in their evaluations. The system relies on detecting Wi-Fi probes sent by mobile phones and received by Wi-Fi monitors installed at different locations. [18] presents a low-cost Raspberry Pi-based Wi-Fi sensing technology to track people and simulate crowds at mass events. [19] estimates crowd densities and pedestrian flows using Wi-Fi and Bluetooth in an airport. [20] monitors pedestrian and cyclists travel-time using Wi-Fi and Bluetooth. Tracking the public

transport usage via Bluetooth and Wi-Fi passive sensing has also been researched in the past for e.g. counting the number of passengers waiting at the bus stops [21].

Several software packages or tools such as Airodump-ng<sup>1</sup>, Tcpdump<sup>2</sup>, Kismet<sup>3</sup>, and Wireshark<sup>4</sup> are also available for capturing Wi-Fi packets. While much work has been done on counting people and crowd in fixed locations, with sensors installed in strategic locations, our approach is aimed at how we can understand occupancy within a dynamic environment with constantly changing geographic locations, with a high noise ratio of passers-by (e.g. people walking nearby, people waiting at bus stops for other lines, passengers in nearby cars, etc.) and able to work with different user types and devices.

### Design of the System

Figure 1 illustrates three main modules in the designed systems. 1) Data collection: sensed Wi-Fi, Bluetooth, and Bluetooth LE timestamped, geotagged sensing are collected and sent to the server. 2) Data processing: the server aggregates data, removes noise (e.g. passers-by, duplicates, etc.) creating a clean segment based head count for each bus journey 3) Data visualisation: final data and estimated passenger counts are presented decision makers via large scale data visualization methodologies.

1 <https://www.aircrack-ng.org/doku.php?id=airodump-ng>

2 <http://www.tcpdump.org/>

3 <https://www.kismetwireless.net/>

4 <https://www.wireshark.org/>

### *Hardware*

Given the requirement of sensors to be deployed in hundreds of locations, the Raspberry Pi (a single-board Linux computer using an ARM-based system-on-a-chip) is an ideal option. In order to sense Wi-Fi/Bluetooth devices, a series of components are integrated with the Raspberry Pi as listed below:

- Processor: a 900MHz quad-core ARM Cortex-A7 CPU
- Wi-Fi Adapter: Ralink RT5370 chipset
- Bluetooth Adapter: CSR Bluetooth 4.0
- SD card: 16GB SD card
- GPS Module: Adafruit Ultimate GPS Breakout – 66 channel w/10 Hz updates

The hardware configuration has low energy consumption (5V 1A power supply), and is significantly lower cost (under £100 including all components). Waterproof cases are used for securing sensors and protecting them from adverse weather conditions.

### *Software*

We use Airodump-ng<sup>5</sup> to enable the Wi-Fi monitor mode for Raspberry Pis, which provides timestamped MAC address of sensed Wi-Fi, number of the frames, and Received Signal Strength Indicator (RSSI). The collected information, along with the RSSI is analyzed to determine the number of the unique sensed devices. While there is no direct means of linking users' personal information and MAC addresses of their Wi-Fi and Bluetooth devices, it is conceivable that user mobility information and habits could be studied to uniquely identify individuals. As a result, storing timestamped

<sup>5</sup> <https://www.aircrack-ng.org/doku.php?id=airodump-ng>

and geotagged records of MAC addresses is highly sensitive and a privacy risk. To address this risk, all the collected MAC addresses from the sensors are transformed via a hash function for privacy issues.

### **Data Processing**

It should be noted that the Wi-Fi/Bluetooth sensing is not a highly accurate process and hence, not designed to provide absolute numbers of passengers. As a result of the design itself, Wi-Fi and Bluetooth sensing under- or over-estimates real counts of people. For example, while monitoring passengers waiting at a bus stop, people who carry Wi-Fi/Bluetooth devices passing by will add noise to the measurement. The relation between the actual number of the crowd/passengers and the monitored number is also not known. Whilst passengers carrying multiple devices may have their Wi-Fi/Bluetooth turned on, there are commuters who do not carry any of these devices or keep the Wi-Fi/Bluetooth off. As mentioned earlier, our approach does not attempt to count distinct passengers, but rather aims at understanding if a reasonable estimate of public transport occupancy can be achieved using such means. In order to overcome the aforementioned issues related to passer by commuters and to improve accuracy, we employ a RSSI and temporal filtering approach (more details below). We validate our results by deploying the sensors on bus journeys and comparing them against ground truth collected manually by counting passengers during 27 trips over five days in a UK metropolitan area.

### *Data Aggregation & Filtering*

In order to remove noise, we employ a multi-pass aggregation and filtering technique to address the following issues: (i) duplicates, (ii) pedestrians passes-

by (iii) passengers waiting at bus stops (iv) people in nearby cars or buildings.

Every time a public means of transport stops or slows down (e.g. at bus stops, traffic lights, traffic jams, etc.) there is an increased likelihood of noise being picked up from e.g. pedestrians walking by or passengers of nearby cars or even buses (particularly in the center of cities). In order to address such cases, we compute factors such as distance from sensor and presence over transport segments (e.g. presence over time and in particular between bus stops) as primary means to filter such noise. RSSI values beyond -90 dB are filtered out to only consider devices in the vicinity of the sensors, while detected duration < 1 min are filtered out. The threshold values are selected based on the test of relationship between the signal strength and distance as shown in Figure 2. The threshold value can vary for different use cases or by using different Wi-Fi adapters.

We have designed a formula to estimate passenger counts. It is assumed that after filtering, the sensed number is smaller than the ground truth since some passengers may not own a smartphone or turned their Wi-Fi on. Therefore, a scale factor is calculated to compensate the filtered signals. The accurate calculation of the scale factor should be based on historical database or a large number of experimental datasets.

- Calculate scale factor:

$$\alpha = \frac{\sum_{i=1}^N \frac{Ground\ Truth_i}{Filtered\ Signal_i}}{N}$$

- Final estimate of passenger counts:

$$Final\ Estimate_i = \alpha * Filtered\ Signal_i$$

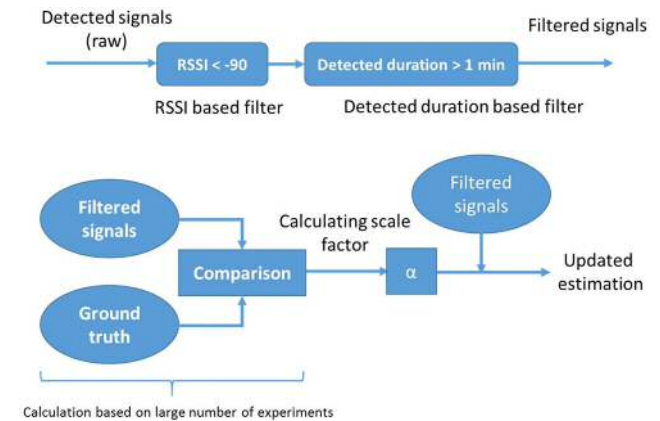


Figure 2: Passenger counting algorithm

## Evaluations

### Evaluation Setting

A set of preliminary evaluations was designed to explore the feasibility and accuracy of the developed sensing system in different public transportation modes for example single-decker bus, double-decker bus and tram as well as varied scenarios, such as different levels of population, residential/commercial areas, different routes etc. The evaluation was conducted over a period of five days, with the experimenter making different trips along with the sensors. Ground truth was collected through manual counting of the number of the current passengers and the number of the passengers getting on or off the bus at every bus stop. The bus journeys were carefully chosen to cover a variety of scenarios. The details of the No. of journeys and the journey durations are shown in Table 1.

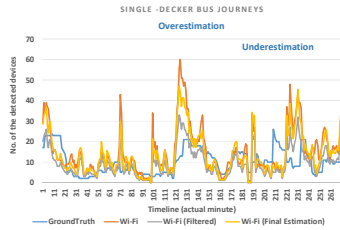


Figure 3: Comparison between Initial sensing results and ground truth

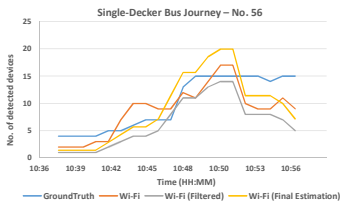


Figure 4: Results comparison between the raw sensed data and filtered data and the ground truth.

| Journey Type      | No. of journeys | Total Journey Duration (mins) | Setting   |
|-------------------|-----------------|-------------------------------|---|
| Single-Decker bus | 8               | 270                           | Trip from a dense city center area to lower density suburbs   |
| Double-Decker bus | 16              | 370                           | Both trips from a dense city center area to different lower density suburbs and trips within dense city areas. Sensors were either on either upper or lower deck. |
| Tram              | 3               | 86                            | Trip from a dense city center to a shopping mall four miles outside the city center   |

**Table 1:** Different transport journeys were conducted to capture data, covering a variety of settings

#### Data Collected

In the five days of the evaluation, in total we collected 6574 unique Wi-Fi MAC address, 444 unique Bluetooth MAC address, and 2259 unique Bluetooth LE MAC address. Total number of bus journeys tracked was 27, with 33% of trips occurring between 10AM and 12PM (average of 93 devices tracked per journey), 11% between 12PM and 2PM (average of 133 devices tracked per journey), 30% between 2PM and 4PM (average of 173 devices tracked per journey), 26% between 4PM and 6PM (average of 226 devices tracked per journey).

#### Analysis

Data analysis was conducted in two steps. An initial manual observation aligning device counts with observer notes was conducted. This was primarily to verify if enough sensing data had been collected and our sensors had provided consistent data. The next process involved aligning manual observations with geotagged timestamped device counts from the sensors. Following which, the raw data and manual observations were compared.

#### Experiment Results

Figure 3 presents a plot of the comparison between the raw data and the ground truth on total single-decker bus journeys. As expected, several examples can be identified where the sensor system over- and underestimate number of passengers. Overestimation can be explained as the noise from the pedestrians who have carried devices on the streets or from the mobile devices in the buildings close to the streets. A filtering technique (as discussed earlier) was developed to remove such overestimated values. Underestimation occurred when there are passengers who do not carry their mobile devices or have turned their Wi-Fi off. This can be corrected by applying a scale factor on the filtered data (as discussed earlier). Figure 4 presents the results achieved after running the filtering and estimation algorithms. As can be observed, the updated sensing result shows a good correlation with the Ground truth (Pearson correlation  $r=0.839$ ,  $p$ -value  $< 0.01$ ).

While in general, the filtering mechanisms appear to have reduced noise significantly, performance was observed to vary between different locations of the city. For example, Figure 5 aligns the geolocation of the

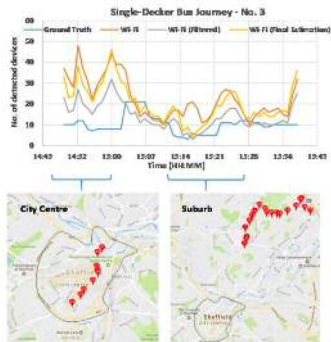


Figure 5: Overestimation in the city center and underestimation in the suburb area.

observations with the number of detected devices. In the city center areas (left), we observe an over-estimation even after filtering. We believe this may due to traffic congestion or slow moving traffic in the city center where cars and public transport may travel at the same speed, often also matching pedestrian speed.

It can be seen from our initial results, the number of the sensed Bluetooth LE devices is about 1/3 of the sensed Wi-Fi devices, which shows a possibility in estimating the passenger counts by sensed Bluetooth LE devices. However, the sensed data of Bluetooth LE devices is sparse, and most of the sensed devices only appear for a few seconds. On account of the sensing reliability, passengers counting estimation by using Bluetooth LE sensing is not considered in this study. It will be further investigated in future.

## Discussion

Although the estimation of passenger load has been improved after applying a generic method for noise reduction and final estimation, there is still observed overestimation in some results (e.g. results presented in Figure 5). The estimation can be further improved with more information (e.g. demographics and geolocation) considered and technologies incorporated.

### *Improve estimation accuracy based on demographics*

It should be noted that the public transport occupancy are influenced by many demographic factors and economic characteristics. For example, after the rush hours in the morning, more elderly people are using the bus services, and they tend not to own a smartphone. A scale factor based on demographics (for example, 60% of the elderly people own a smartphone) can help improve the estimation accuracy in this situation.

### *Reduce noise caused by traffic jam*

As the results presented in Figure 5, traffic jam is causing significant over-estimation even after the filtering approach is applied. It dues to the fact that noises from other vehicles or passers-by have been sensed. In order to solve this issue, certain criteria can be added to the filter. If the device left the bus not at a bus stop, it can be treated as the noise in a traffic jam. Also in order to determine whether the bus is experiencing a traffic jam, a GPS module or an activity tracker (e.g. accelerometer) can be used to identify it.

### *Reduce the noise at bus stop*

At rush hours, buses usually stop for longer time at bus stops for loading passengers. The sensors can pick up noises from the passenger who wait for other bus services. Therefore, identifying the bus stop and knowing how long the bus is stopped for the stop will help to eliminate this type of noise. In order to accurately identify the bus stop and how long the bus stopped at bus stop, a series of sensing techniques and information is needed, for example, the bus routes information from the Bus Company, GPS module, and activity tracker. The sensed device only appeared in the period when bus stopped at the bus stop would be identified as the passengers who were waiting for other bus services. Furthermore, these sensing results can be cross checked with the sensor at the bus stop (if there were).

## Conclusion and Future Work

We have evaluated our monitoring system for five days experiment on bus journeys. Initial results are promising and show good correlation with ground truth after filtering out unwanted signals and application of a scale factor.



Figure 6: Smartphone based mobile sensor

The observed variance based on locations pose a challenge to estimate passenger counts. Future work will address this in two major strands: first, we plan to exploit contextual information regarding the journeys – the filtering and estimation algorithms will consider the locations (e.g. POIs, city center, bus station areas) as well as time of day (e.g. office hours, lunch time) to automatically define thresholds. We expect this to significantly improve the passenger estimation. The next activity will focus on exploiting background knowledge on bus trips, timings and passenger types. For example, morning commuters are typically office goers, likely to be sensed via smartphones. However, many afternoon passengers are elderly citizens, likely not to possess smartphones and hence may not be sensed. Future work will also involve a large scale evaluation in three cities (Birmingham, Santander and Turin) as a part of the Seta project.

In future, we will also evaluate the mobile version of our system. We have also developed a mobile version of sensors (see Figure 6). The mobile sensors share the same working principle as the Raspberry Pi based static sensors. The mobiles can be made available to a bus driver or a citizen who carries it during their bus journeys. The sensor can also be configured to be associated to public vehicles, e.g. buses, taxis, ambulances, hire-bicycles. The use of a mobile phone immensely simplifies the process of sensing – users can carry such sensors while conducting their daily activities throughout cities, using public transport and reports can be automatically generated for processing and analysis. Albeit the simplicity in this design, the process of developing the sensors is not straightforward – only a few older phones allow their Wi-Fi chipset to be turned into monitor mode but it can be made for

certain firmware (e.g. HTC Desire). In addition, it can be potentially cheaper solution as available secondhand is less than £30. Also have the advantage of a screen and mobile network communication.

### Acknowledgement

This work is supported by Seta (A ubiquitous data and service ecosystem for better metropolitan mobility) project funded by the European Commission as part of the Horizon 2020 programme under contract 688082.

### References

1. Givoni, M. and Banister, D. eds., 2013. Moving towards low carbon mobility. Edward Elgar Publishing.
2. Hickman, R., Hall, P. and Banister, D., 2013. Planning more for sustainable mobility. *Journal of Transport Geography*, 33: 210-219.
3. Ceder, A., 1984. Bus frequency determination using passenger count data. *Transportation Research Part A: General*, 18(5-6): 439-453.
4. Shen, X., Feng, S., Li, Z. and Hu, B., 2016. Analysis of bus passenger comfort perception based on passenger load factor and in-vehicle time. *SpringerPlus*, 5(1): 1-10.
5. Eboli, L. and Mazzulla, G., 2007. Service quality attributes affecting customer satisfaction for bus transit. *Journal of public transportation*, 10(3): 2.
6. Tirachini, A., Hensher, D.A. and Rose, J.M., 2013. Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. *Transportation research part A: policy and practice*, 53: 36-52.
7. Chan, A.B. and Vasconcelos, N., 2012. Counting people with low-level features and Bayesian regression. *IEEE Transactions on Image Processing*, 21(4): 2160-2177.



8. Fujisawa, S., Hasegawa, G., Taniguchi, Y. and Nakano, H., 2013. Pedestrian counting in video sequences based on optical flow clustering. *International Journal of Image Processing*, 7(1):1-16.
9. Bernini, N., Bombini, L., Buzzoni, M., Cerri, P. and Grisleri, P., 2014, September. An embedded system for counting passengers in public transportation vehicles. In *Mechatronic and Embedded Systems and Applications (MESA), IEEE /ASME 10th International Conference on* (pp. 1-6).
10. Yan, L., Chao, L., Wei, P.K. and Tao, S., 2015. Crowd Counting Using Wireless Infrared Distance Sensors for Indoor Environments. In *Informatics and Communication Technologies for Societal Development*, 1-8. Springer India.
11. Oberli, C., Torres-Torriti, M. and Landau, D., 2010. Performance evaluation of UHF RFID technologies for real-time passenger recognition in intelligent public transportation systems. *IEEE transactions on intelligent transportation systems*, 11(3):748-753.
12. Trépanier, M., Tranchant, N. and Chapleau, R., 2007. Individual trip destination estimation in a transit smart card automated fare collection system. *Journal of Intelligent Transportation Systems*, 11(1): 1-14.
13. Handte, M., Iqbal, M.U., Wagner, S., Apolinarski, W., Marrón, P.J., Navarro, E.M.M., Martinez, S., Barthelemy, S.I. and Fernández, M.G., 2014, March. Crowd Density Estimation for Public Transport Vehicles. In *EDBT/ICDT Workshops* (pp. 315-322).
14. Kostakos, V., Camacho, T. and Mantero, C., 2013. Towards proximity-based passenger sensing on public transport buses. *Personal and ubiquitous computing*, 17(8): 1807-1816.
15. Stange, H., Liebig, T., Hecker, D., Andrienko, G. and Andrienko, N., 2011, November. Analytical workflow of monitoring human mobility in big event settings using bluetooth. In *Proceedings of the 3rd ACM SIGSPATIAL international workshop on indoor spatial awareness* (pp. 51-58).
16. Weppner, J. and Lukowicz, P., 2013, March. Bluetooth based collaborative crowd density estimation with mobile phones. In *Pervasive computing and communications (PerCom), 2013 IEEE international conference on* (pp. 193-200).
17. Musa, A.B.M. and Eriksson, J., 2012, November. Tracking unmodified smartphones using wi-fi monitors. In *Proceedings of the 10th ACM conference on embedded network sensor systems* (pp. 281-294).
18. Bonne, B., Barzan, A., Quax, P. and Lamotte, W., 2013, June. WiFiPi: Involuntary tracking of visitors at mass events. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a* (pp. 1-6).
19. Schauer, L., Werner, M. and Marcus, P., 2014, December. Estimating crowd densities and pedestrian flows using wi-fi and bluetooth. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services* (pp. 171-177). ICST.
20. Abedi, N., Bhaskar, A., Chung, E. and Miska, M., 2015. Assessment of antenna characteristic effects on pedestrian and cyclists travel-time estimation based on Bluetooth and WiFi MAC addresses. *Transportation Research Part C: Emerging Technologies*, 60: 124-141.
21. Oransirikul, T., Nishide, R., Piumarta, I. and Takada, H., 2014. Measuring Bus Passenger Load by Monitoring Wi-Fi Transmissions from Mobile Devices. *Procedia Technology*, 18: 120-125.