Estimating Pedestrian Densities, Wait Times and Flows using Wi-Fi and Bluetooth Sensors

Abdullah Kurkcu, Kaan Ozbay

This is the author's version of a work that was published in Transportation Research Record: Journal of the Transportation Research Board in January 2017. The final version can be found here: https://doi.org/10.3141/2644-09

Estimating Pedestrian Densities, Wait Times and Flows using Wi-Fi and Bluetooth Sensors

Abdullah Kurkcu, M.Eng. (Corresponding Author)
Graduate Research Assistant, Department of Civil & Urban Engineering,
UrbanITS Center,
Tandon School of Engineering,
Center for Urban Science + Progress (CUSP), New York University (NYU)
Six MetroTech Center, 4th Floor,
Brooklyn, NY 11201, USA
Tel: 1-(646)-997-0538
Email: ak4728@nyu.edu

Kaan Ozbay, Ph.D.
Professor, Department of Civil & Urban Engineering,
UrbanITS Center,
Tandon School of Engineering,
Center for Urban Science + Progress (CUSP), New York University (NYU)
Six MetroTech Center, 4th Floor,
Brooklyn, NY 11201, USA
Tel: +1 (646) 997 0552
Email: kaan.ozbay@nyu.edu

Word count: 4644 Text + 11 tables/figures x 250 words (each) = 7394 Words

TRR Paper number: 17-02095

Submission Date: August 1, 2016

ABSTRACT

Monitoring non-motorized traffic is gaining more attention in the context of transportation studies. Most of the traditional pedestrian monitoring technologies focus on counting pedestrians passing through a fixed location in the network. It is thus not possible to anonymously track the movement of individuals or groups as they move outside of each particular sensor's range. Moreover, most agencies do not have continuous pedestrian counts mainly because of technological limitations. However, wireless data collection technologies can capture crowd dynamics by scanning mobile devices. Data collection that takes advantage of mobile devices has gained much interest in the transportation literature due to its low cost, ease of implementation and richness of captured data (*1*). In this paper, algorithms to filter and aggregate data collected by wireless sensors and how to fuse additional data sources to improve the estimation of various pedestrian based performance measures are investigated. Procedures to accurately filter the noise in the collected data and find pedestrian flows, wait times, and counts using wireless sensors are presented. The developed methodologies are applied to a 2-month long public transportation terminal data collected by six sensors. The results pointed out that if the penetration rate of discoverable devices is known, it is possible to accurately estimate the number of pedestrians, pedestrian flows and average wait times within the detection zone of the developed sensors.

INTRODUCTION & MOTIVATION

Human movement behavior research has received increasing attention particularly in the field of transportation planning. The traditional methods for pedestrian mobility monitoring include surveys, fixed pedestrian counters, and vision-based technologies. However, these techniques are neither easy to implement nor cost-effective. In addition, video-based technologies rely on a clear view of the crowd over a limited spatial range which requires integration of data from a number of cameras over the whole spatial range for which mobility data is collected. In recent years, several studies have been reported in the literature on automating pedestrian detection or counting to explore economic and reliable methods (*2*-*5*). These researchers reviewed the available automated pedestrian counting technologies such as infrared and thermal sensors (*6*). With the increase in smart devices, research has started focusing on tracking mobile phones to estimate pedestrian movements. If the detection system is equipped with Wi-Fi and Bluetooth receivers, it is possible to capture Origin-Destination (OD), travel time, wait time and flow information for some subset of the pedestrians with visible Wi-Fi and Bluetooth devices. People with electronic devices, such as most cell phones, tablets, and computers carry unique information, a Media Access Control (MAC) address, in their devices that can be used to collect pedestrian data for estimating measures such as travel time (*7*). This type of traffic detection systems can be supplemented by traditional sensing technologies to improve crowd monitoring systems (*8*).

MAC addresses are the most common unique identifiers in IEEE 802 network technologies. There are 6 bytes/48 bits, making it possible to generate 2^{48} potential unique MAC addresses. The first three bytes contain an organizationally unique identifier (OUI), and the following three are assigned by the organization in any manner as long as it is unique. Every Bluetooth or Wi-Fi device is defined by a MAC address. Therefore, individual devices can be tracked, and this feature has been utilized in various applications and data collection processes in the literature.

To be able to manage transportation systems efficiently, information about non-motorized traffic is required. However, decision makers and transportation officials in the U.S. have not yet extensively examined non-motorized traffic (*9*). In addition, most agencies lack comprehensive pedestrian counts mainly because of technological limitations. Some of these challenges can be explained as follows:

- Unlike motorized vehicles, pedestrians do not travel in fixed lanes or paths and make unpredictable movements.
- Pedestrians sometimes travel very closely to each other creating platoons, and some sensors have difficulty counting individuals within the group (3).
- The number of locations for which pedestrian data are needed is exponentially higher than is the number needed for monitoring vehicular traffic.

Although FHWA's Traffic Monitoring Guide (TMG) does not address technologies such as Wi-Fi and Bluetooth sensing for O-D or travel time, it summarizes the potential options for pedestrian counting technologies and respective costs (*10*). According to the FHWA's guide, calibration and validation procedures should be implemented to ensure that pedestrian counts are within the bounds of acceptable accuracy. TABLE 1 below illustrates available technologies and costs adapted from that guide; the table is updated with the addition of relevant information about wireless sensors. The proposed sensors cost approximately $100, including additional parts such as Wi-Fi and Bluetooth USB antennas.

Most of the common pedestrian monitoring technologies mentioned in the table focus on counting pedestrians according to the location specific point data in the network. However, it is not possible to anonymously track pedestrians or groups as they move outside of each particular sensor's range. Tracking pedestrians may be achieved by using video imaging and matching the same pedestrians but not preferred due to the high cost and computational complexity. Therefore, there is a need for building low-cost, customized ubiquitous sensors. The sensors and the codes used to collect wireless traces were developed in-house in this study. These sensors use wireless technologies and algorithms to track individual movements, measure wait times, and estimate flows. One disadvantage of using such sensors is that not everyone carries a detectable smart device, and some carry more than one. Estimating flows and counts rely on some assumptions about the crowd and conditional factors, depending on the site of the study.

In this study, methods on how to anonymize, filter and aggregate traces of pedestrians, and how to fuse additional data sources to enhance existing filtering and counting algorithms are discussed. The proposed methods are evaluated using various data collection scenarios. After the evaluation period, developed techniques are applied to a 2-month long collection of public transportation terminal data and the results are reported.

BACKGROUND

Bluetooth technology is the global wireless standard enabling the communication between smart devices using radio transmissions. The key properties of Bluetooth are robustness, low power, and low cost (*11*). Although not all Bluetooth devices are discoverable, in general, it has been reported that 5%-12% of devices are discoverable via Bluetooth (*12*). When the user is looking for other devices to connect to using Bluetooth, the device will be active and visible to all other devices on the network. Developed sensors can detect only those devices that are actively looking for other Bluetooth devices or are already connected to another device via Bluetooth.

Wi-Fi is a technology defined in the IEEE 802.11 standard that allows electronic devices to connect to a wireless local area network with the use of radio bands. Whenever devices try to connect to a wireless local area network, they send out a probe request. A probe request is a special frame that asked for information from either a specific access point or all access points in the area. By sending a probe request, the wireless device is making an active scan of networks. Probe requests contain the MAC address of the sender; the service set identifier, which is the network name; and the received signal strength indication.

To ensure the maximum amount of anonymity, data collection systems that will use signals from Wi-Fi and Bluetooth enabled mobile devices should not store personally identifiable data; therefore, MAC addresses collected by sensors should not be associated with specific users. In some studies in the literature, the electronic identifier of the mobile device of the detected agent is converted into an encrypted hash code. MAC addresses that are not matched are deleted at the site (*7*). In others, it has been stated that privacy concerns for the end users are a nonissue when the data collected through Bluetooth. The MAC addresses were kept anonymous without being directly tied to individuals (*1; 13; 14*). In a 2010 research study by the Texas Transportation Institute (2010), a routine is added to Bluetooth data collection software to encrypt collected MAC addresses. This is done to make sure that actual device addresses are not stored anywhere, but rather a random set of characters is used (*15*). In this paper, a similar technique is proposed for the encryption to ensure maximum privacy while maintaining persistent records. The code that grabs the MAC address from probe requests only checks the last three octets. Instead of keeping all digits of the MAC address, to protect the embedded private information, only the part that contains the last six digits is read. For example, if the MAC address of a device is MM:MM:MM:SS:SS:SS, only SS:SS:SS is kept. After the last six digits are retrieved, that information is then encrypted with an encryption key and stored on the instrument. This technique provides an extra layer of protection. It also keeps the MAC addresses unique for approximately 98% of the detected devices. The encryption key is first randomly generated on a remote server. After the initial key is generated, it is then encrypted again before being uploaded to the devices on site. Moreover, it is proposed that for most of the field applications, individual data points be aggregated and deleted on site depending on the type of data collection. For example, to obviate the need to keep even highly anonymized and encrypted individual data points, sensors will keep the aggregated counts only if they are calibrated for counting.

Earlier efforts found in the transportation literature dealing with the determination of individual movements mostly focused on Bluetooth-based detection. The work presented by Ahmed et al. (*16)* is among the first to utilize Bluetooth detection for vehicle monitoring. The contribution of this work is the deployment of a very low cost and low power device/software combination for transit related OD estimation applications for the first time. Kostakos (*17)* used Bluetooth devices to trace passenger journeys on public buses and derive passenger OD matrices. Bullock et al. (*18)* deployed a Bluetooth tracking system at the new Indianapolis International Airport to measure the time for passengers to transit from the non-sterile side of the airport (pre-security), clear the security screening checkpoint and enter the walkway to the sterile side. Hamedi, Haghani and Sadabadi (*19)* investigated the quality of vehicle probe data using new traffic surveillance devices based on Bluetooth technology. Their results showed that this technology is a promising method for collecting high-quality travel time data that can be employed for evaluating other sources of travel time and intelligent transportation systems (ITS) applications. Haseman, Wasson and Bullock (*20)* also used Bluetooth probe data from multiple field collection sites to quantify delay and to assess diversion rates at a rural Interstate Highway work zone along I-65 in Northwestern Indiana.

Malinovskiy, Saunier and Wang (*7)* presented a study of pedestrian detection using Bluetooth at two separate sites. They investigated the feasibility of using Bluetooth technology for pedestrian studies and found out that it can provide useful information for pedestrian travel behavior. Barceló et al. (*13)* used travel time data captured by Bluetooth sensors to estimate time -dependent OD matrices in simulation tests. Lees-Miller, Wilson and Box (*21)* tried recovering the path of a vehicle by using Bluetooth detection data. The proposed approach was able to reconstruct vehicle trajectories outperforming a simple deterministic strategy by 30-50%. Michau et al. (*22)* showed that the position of the detectors is of great importance, and the wireless signals are easily weakened by physical conditions as well as weather. The detection process of a Bluetooth device can be described as a cycle during which the sniffer will transmit messages on different range of frequencies and waiting for devices to pick up that message. Therefore, it was concluded that Bluetooth devices have to be in a discoverable mode for about 10 seconds within the detection zone in order to detect them (*22*). Laharotte et al. (*14)* provided some insights on how Bluetooth data can be used for vehicular flow forecasting. Their filtering algorithm reconstructs traffic states at a network scale using nonparametric pattern recognition techniques with a k-nearest-neighbors (kNN) procedure. Their prediction of the network traffic state with a kNN approach showed convincing results using 31 days of data.

Integration of Wi-Fi systems with Bluetooth sensors can be found in recent studies dealing with real-time data collection and monitoring of pedestrian networks. Lesani and Miranda-Moreno (*23)* investigated the advantages and the feasibility of a Wi-Fi data collection system as an alternative and a supplement to Bluetooth technology. They found that the detection rate for Bluetooth is as low as 2.0%, and the combination of Wi-Fi and Bluetooth systems showed promising results. Hourly travel time estimation errors were around 3.8%. The average and median prediction errors of pedestrian flows were 15% and 9% respectively. Weppner et al. (*24*) (*25*) used Bluetooth scanners to count the number of devices in a fixed region. Nicolai and Kenn (*26*) presented a method to find out the relationship between detected Bluetooth devices and the ground truth data. Kalogianni et al. (*27*) used passive Wi-Fi scanning method to sense the movements of students, employees and visitors in a university campus. They investigated what kind of patterns can be captured by WiFi monitoring and how people utilize the buildings at the campus. The results pointed out that passive Wi-Fi monitoring is an effective way to identify building usage and movement between buildings. Bonne et al. (*28*) a low-cost crowd counting system based on a single-board computer with the addition of a LED to provide a status indicator and an Android cell phone as an operator. 15 devices were deployed in a music festival and 4 in a university campus. They concluded that tracking visitors at mass events can be achieved by using Raspberry PI sensors at a very low cost. Abedi et al. (*8*) used a commercial sensor with the capability of scanning both Bluetooth and Wi-Fi addresses simultaneously. They compared the standards for both technologies regarding architecture, discovery time, signal strength and popularity of use. The results pointed out that Wi-Fi has shorter discovery time, the distance from the sensor can be estimated based on the signal strength, and Wi-Fi is accepted as the more appropriate standard compared to Bluetooth for pedestrian data collection. Abedi, Bhaskar and Chung (*8)* evaluated antenna characteristics and concluded that the bigger antenna gains capture more data, but they may not be useful for small scales of monitoring due to overlapping detection areas. Schauer et al. (*29*) used both Wi-Fi and Bluetooth sensors to estimate crowd densities and pedestrian flow at a major German airport. Additional studies are dealing with pattern mining in tourist attraction visits (*30*), Bayesian approach to detect destinations (*31*), and location popularity and visit patterns (*32*) can be found in the literature.

Although almost all of these applications used similar datasets, only a few developed their own sensors and comprehensive techniques for removing erroneous detections (*33*). In addition, measures other than pedestrian movement and O-D data were not investigated. In this study, sensors developed by the research team to collect Wi-Fi and Bluetooth traces are used to improve scenario based filtering algorithms and accurately report measures such as pedestrian counts, wait times and time-dependent O-D patterns.

RESEARCH METHOD

It is possible to detect the proximity of personal electronic devices with Wi-Fi and Bluetooth when they are actively looking for other devices. Wireless sensors can not only detect the mobile devices around them but also non-mobile devices, access points and other networks disseminating their presence to the network. Therefore, it is critical to build efficient algorithms to filter the raw data captured by the sensors. In this paper, four filtering algorithms are developed to accurately refine wireless traces. The initial filter will remove the devices that are far away from the sensors and the non-mobile devices in the captured data. After this initial cleaning, pedestrian counts, flows and wait times can be calculated with the, respectively, moving blocks, flow and wait time filters. FIGURE 1 shows the flowchart of the applied algorithms to the collected traces and their results.

Virtual Sensors – Bus Time

A web scraper is also developed to retrieve real-time bus schedules and delays for the corresponding bus lines from the transit authority's web application. It behaves as a Virtual Sensor (*34*) to track bus departure times and delays. This information will be used in the evaluation of wait times to remove the erroneous detections efficiently and to calculate maximum acceptable wait times at the stations.

Pedestrian Sensors

Hardware

Raspberry PI (*35*) is a low cost small sized computer that is compatible with monitors or televisions. It uses a standard keyboard and mouse. Pi supports object-oriented programming languages, such as Scratch and Python. It is capable of processing tasks that one can expect a desktop computer to do. In addition, it is possible to build the Pi so that it will have the ability to interact with the environment. As reported in the literature, the Raspberry PI and similar mini PCs have been used to detect motion (*36*), measure the noise (*37*) and air quality (*38*), and monitor the environment (*39*) as well as for various smart city applications (*40-42*) in the literature. The developed sensor can be seen in FIGURE 2.

Software

Installation of Raspbian, Linux operating system, is minimal and the system comes with the PI's SD card installed. The Aircrack-ng library is used to create an access point and sniff the network, and SQLite3 is used as a database to store the data. The signal strength (received signal strength indication) of the devices can be used to create a detection zone. With tuning, a circular detection circle is created that when crossed, can trigger the detection system to store the information in SQLite tables.

Data Filtering Algorithms

Seven trial field tests with varying lengths, from a couple of hours to a couple of weeks, are conducted before an actual field test to develop and improve data filtering algorithms. To evaluate and supplement sensor's crowd count, manual counts that are done to be collected as ground truth data are also reported. TABLE 2 summarizes the conducted field tests. Parameters and inputs to these algorithms need to be tweaked, depending on the purpose and the location of the study.

Initial Filtering of Data

The initial filtering process starts with finding the addresses of the devices that occurred most often. If a device occurs in the database more frequent than every 10 minutes on average for a 6-h period, it is removed. The reason behind this initial filtering is to eliminate non-mobile devices captured in the database, devices of the staff, and any other smart device that has not moved during the test time period. The second filter is applied only to the sensors at the selected bus gates to detect devices with signal strengths that are lower than some threshold value. If a device is detected at least once with an RSSI value stronger than -80 dBi, it is retained in the database; otherwise, it is removed. The selection of -80 dBi corresponds approximately to a 15-meter radius proximity to the sensor.

Moving Blocks - Counting Crowds

A new filtering algorithm is designed to count discoverable people who are within the detection range of the sensor. It aims to filter pedestrians moving in circles, or going back and forth. The goal here is to address the problem of double counting due to such movements. The algorithm creates a cycle block for the first 5 minutes and stores every detected address in it. It then checks if the detected MAC addresses in the 6^{th} minute can be found in the first 5 minutes. If a common MAC address is detected, it is removed from the existing minute's count. For the next time period, it creates another 5-minute block starting from the 2nd minute to 6th minute and checks the mac addresses detected in the 7th minute with this new 5-minute block. The visualization of this method can be seen in FIGURE 3a. The moving block algorithm successfully removed a significant amount of noise in the count data during the field tests. However, it is prone to errors at the beginning and the end of the study due to the limited block availability. The counts become more reliable with a 5-minute warm-up and a cool down period. It should be noted that the algorithm's performance to accurately count pedestrians is heavily dependent on the ratio of people with discoverable devices.

Estimation of Pedestrian Flows

Flow data can be captured by finding overlapping anonymized and encrypted MAC addresses between two sensors. FIGURE 3b shows the flow estimation algorithm. For each record, the time it took for an individual to go to one sensor from another is calculated from the data. The distances between the sensors are known and irregular travel times can be filtered out by using the walking speed. The data points are removed if the calculated speed does not fall within the 95th percentile confidence interval. It is assumed that pedestrians have large inter-individual differences and the desired speed of an individual following a Gaussian distribution with mean 1.2 m/s and standard deviation 0.3 m/s (***43***).

Wait Time Calculation

The time difference between the first and the last occurrence of anonymized and encrypted MAC addresses are stored in a temporary table. It is important to emphasize the fact that this data does not need to be saved and can be deleted upon the completion of the calculations for 5 minute or 15 minute periods. FIGURE 3c visualizes the proposed filtering algorithm. Individual pedestrian wait times are calculated for each hour, and they are assigned to hour slices depending on the first detection time. Using the real-time bus data, the longest headway between two buses is found for respective gates. Therefore, pedestrian wait times that are longer than that value and shorter than 30 seconds are removed from the temporary table for further analysis.

EXPERIMENTAL RESULTS OF FILTERING ALGORITHMS

Sensors are deployed in six different locations at a transit terminal. Two are placed at the main entrances and four are located at the busiest gates. The data stored on sensors collected every week for the first month and every two weeks for the following month. One week in which all the sensors worked without a problem is selected for further investigation. Developed filtering algorithms are applied to the collected data and results are reported in this section.

Initial Filtering

Initial filtering algorithm described above is applied to the week-long data to capture the patterns of the number of detected devices among different days. FIGURE 4a shows the results from Monday to Sunday for the test week. It can be clearly seen that the number of detected devices follow a decreasing trend during the week. The busiest times at this specific entrance point are experienced on Monday. The evening peak period on Friday is flatter stretching over a longer time period approximately from 3 PM to 9 PM. There are commercial stores nearby Entrance A, thus, FIGURE 4a illustrates not only the number of passengers entering the terminal but also customers visiting these commercial facilities. Detected Devices for a Day

FIGURE 4b shows the number of detected devices for Wednesday. The busiest hour is between 7 AM and 8 AM with 2126 detected devices. The patterns of the device detection more or less remain the same among different weekdays except for Sunday. This can be explained by buses having less frequent service times and lower number of riders and some commercial stores being closed on Sundays at the terminal.

Moving Block Algorithm Results

Two indoor field tests with high foot traffic are conducted to test the moving block algorithm, and FIGURE 5 shows the results. An observer counted pedestrians present by the sensor for at least an hour at 5-minute intervals for each test. Although manual counts are accepted as ground truth, it should be noted that these counts are subject to the usual human error especially when the count time is noted. This can be improved by having a video feed of the study location providing more accurate pedestrian counts. In addition, field tests were conducted in a campus environment and the penetration rate of discoverable devices was assumed to be 100%. The results pointed out that the accuracy of the total count is 97% in the first test and 92% in the second test. However, it failed to capture the variation in counts after 2:30 PM in the second field test as it can be seen from FIGURE 5. One potential explanation for this difference can be devices with wireless features turned off. It is also possible that some of the pedestrians may be double counted by the sensor. The variables that the algorithm uses remained the same in both cases.

The same code is applied to the data collected at Entrance A of the terminal test location to evaluate the pedestrian traffic at a fixed location. FIGURE 6 shows the pedestrian count results. Clearly, the distribution of the number of detected devices is similar to the pedestrian count data. However, the count data is different than simply the number of devices detected at the scene because the algorithm eliminates the double counts in the dataset by consistently checking the reoccurring devices throughout the moving block.

Pedestrian Flows

shows the pedestrian flows from entrances to the gates. While some of the gates such as Gate 1 and Gate 2 are heavily used by passengers that enter through respectively Entrance B and A, other gates most likely have passengers originating from other entrances. In addition, Entrances A and B are closer to Gate 1 and 2 than other gates. There are 37,240 distinct mobile devices detected at Gate 2. Pedestrians using Entrances A and B constitute 26% of all foot traffic at this gate.

Time-dependent O-D pairs can also be created using the wireless sensor data. FIGURE 7 shows the temporal distribution of the detected pedestrian intensity traveling from Entrance A to Gate 2 for a week. The gate experienced its busiest period on Tuesday between 4 PM and 8 PM as it can be seen in the figure below.

Wait Times

Wait times at the gates are studied by comparing them with actual bus schedules by the code described in the previous section. FIGURE 8 visualizes wait times for each detected device. The horizontal bars reflect the wait times and the vertical lines in red and green show the scheduled and actual bus departures respectively. The actual bus departure information is collected by using the Virtual Sensor algorithm explained in the methodology section of this paper. Departure delays up to 5 minutes are experienced at this gate. Buses also may arrive at their gate earlier than the departure time and open their doors for boarding. Therefore, it is more accurate to create a 5-min buffer zone around the scheduled departure times. Even though successive filters are applied to the data, there are still some outliers which may decrease the accuracy of the calculated wait times. For example, the detected devices p1001 and p1004 leave the platform right after the bus departs around 8 AM. These data points should be excluded to calculate wait times more accurately.

CONCLUSIONS AND FUTURE WORK

Algorithms developed to filter and aggregate the data collected by Wi-Fi/Bluetooth sensors and how to fuse an additional data source to improve the estimation of wait times are presented in this paper. The developed methodologies are applied to a 2-month long data collected in a transit terminal. The sensors are located at the two main entrances as well as 4 passenger gates. There are of course limitations of these procedures that deserve mention. Short living network addresses, non-mobile devices that transmit intermittent probe requests and devices that are detectable at a low frequency can reduce the accuracy of the developed algorithms. However, the main contribution here is to alleviate the inaccuracies originated from the noise that are inherent in the collected wireless traces. All of the proposed algorithms aim to remove low-quality detections, eliminate periodic/cyclic behavior, and improve detection and counting performance of devices.

The initial filtering of the data showed that it is probable to capture re-occurring patterns of the passengers in the terminal. The peak periods and busiest hours can also be detected at sensor locations. This information makes it easier to estimate passenger demand at a transit terminal. However, the results from the initial filtering algorithm only represent the number of detected devices and should not be used as actual pedestrian counts.

The moving block algorithm can provide more accurate representation of the pedestrian counts at fixed sensor locations because it can efficiently eliminate the double counts and non-mobile devices. The results from field tests showed that the moving blocks algorithm can reach approximately 90% accuracy with the availability of data 5 minutes before and after the study period if the discoverability rates of devices are close to 100%. If this location and time specific penetration rate is known for other study sites, the moving block filter can be used to factor up the Bluetooth and Wi-Fi samples to account for total pedestrian counts. The flows between sensors indicated that some certain entrances are more heavily used to reach specific gates. Although pedestrian flows provide an indicator of how the entrances are utilized, it should be noted that there are more than 2 entrances and 4 gates at this terminal. Therefore, more sensors should be installed in the future to locate the most heavily used entrances and corresponding gates. In conclusion, we suggest that the wireless data should be used with great care and the well-tested filters have to be used to clean the collected data. In a future study, these results based on one week should be supplemented with additional long-term data to ensure that they represent seasonal variations.

ACKNOWLEDGMENTS AND DISCLAIMER

The contents of this paper only reflect views of the authors who are responsible for the facts and accuracy of the data and results presented herein. The contents of the paper do not necessarily reflect the official views or policies of sponsoring agencies.

REFERENCES

[1] Carpenter, C., M. Fowler, and T. Adler. Generating Route-Specific Origin-Destination Tables Using Bluetooth Technology. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2308, 2012, pp. 96-102.
[2] Bu, F., R. Greene-Roesel, M. C. Diogenes, and D. R. Ragland. Estimating pedestrian accident exposure: automated pedestrian counting devices report. *Safe Transportation Research & Education Center*, 2007.
[3] Ozbay, K., B. Bartin, H. Yang, R. Walla, and R. Williams. Automatic pedestrian counter.In, 2010.
[4] Yang, H., K. Ozbay, and B. Bartin. Enhancing the quality of infrared-based automatic pedestrian sensor data by nonparametric statistical method. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2264, 2011, pp. 11-17.
[5] ---. Investigating the performance of automatic counting sensors for pedestrian traffic data collection.In *Proceedings of the 12th World Conference on Transport Research, Lisbon, Portugal, No. 1115*, 2010.
[6] Ozbay, K., H. Yang, and B. O. Bartin. Calibration of an infrared-based automatic counting system for pedestrian traffic flow data collection.In *Transportation Research Board 89th Annual Meeting*, 2010.
[7] Malinovskiy, Y., N. Saunier, and Y. Wang. Analysis of pedestrian travel with static bluetooth sensors. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2299, 2012, pp. 137-149.
[8] Abedi, N., A. Bhaskar, and E. Chung. Bluetooth and Wi-Fi MAC address based crowd data collection and monitoring: benefits, challenges and enhancement.In *Australasian Transport Research Forum (ATRF), 36th, 2013, Brisbane, Queensland, Australia*, 2013.
[9] Lindsey, G., K. Nordback, and M. Figliozzi. Institutionalizing Bicycle and Pedestrian Monitoring Programs in Three States. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2443, 2014, pp. 134-142.
[10] USDOT, and FHWA. Traffic Monitoring Guide
In *Federal Highway Administration Traffic Monitoring Guide*, September 2013.
[11] Jie, S., L. Na-na, C. Ji-lin, D. Yong-feng, and Z. Zheng. Design and implementation of intelligent transportation system based on GPRS and Bluetooth hybrid model.In *Automation and Logistics, 2008. ICAL 2008. IEEE International Conference on*, IEEE, 2008. pp. 1381-1385.
[12] Brennan Jr, T. M., J. M. Ernst, C. M. Day, D. M. Bullock, J. V. Krogmeier, and M. Martchouk. Influence of vertical sensor placement on data collection efficiency from bluetooth MAC address collection devices. *Journal of Transportation Engineering*, 2010.
[13] Barceló, J., L. Montero, M. Bullejos, O. Serch, and C. Carmona. A Kalman filter approach for the estimation of time dependent OD matrices exploiting Bluetooth traffic data collection.In *Transportation Research Board 91st Annual Meeting*, 2012.
[14] Laharotte, P.-A., R. Billot, E. Faouzi, and H. A. Rakha. Network-Wide Traffic State Prediction Using Bluetooth Data.In *Transportation Research Board 94th Annual Meeting*, 2015.
[15] Puckett, D. D., and M. J. Vickich. Bluetooth®-based travel time/speed measuring systems development.In, 2010.
[16] Ahmed, H., M. EL-Darieby, B. Abdulhai, and Y. Morgan. Bluetooth-and Wi-Fi-based mesh network platform for traffic monitoring.In *Transportation Research Board 87th Annual Meeting*, 2008.
[17] Kostakos, V. Using Bluetooth to capture passenger trips on public transport buses. *arXiv*

preprint arXiv:0806.0874, 2008.
[18] Bullock, D., R. Haseman, J. Wasson, and R. Spitler. Anonymous bluetooth probes for measuring airport security screening passage time: the indianapolis pilot deployment.In *Transportation Research Board 89th Annual Meeting. CD-ROM. Transportation Research Board, Washington DC*, 2010.
[19] Hamedi, M., A. Haghani, and F. Sadabadi. Using Bluetooth technology for validating vehicle probe data.In *16th ITS World Congress and Exhibition on Intelligent Transport Systems and Services*, 2009.
[20] Haseman, R. J., J. S. Wasson, and D. M. Bullock. Real time measurement of work zone travel time delay and evaluation metrics using bluetooth probe tracking. *TRB Paper*, 2009, pp. 10-1442.
[21] Lees-Miller, J., R. E. Wilson, and S. Box. Hidden markov models for vehicle tracking with bluetooth. 2013.
[22] Michau, G., A. Nantes, E. Chung, P. Abry, and P. Borgnat. Retrieving dynamic origin-destination matrices from Bluetooth data. 2014.
[23] Lesani, A., and L. F. Miranda-Moreno. Development and Testing of a Real-Time Wifi-Bluetooth System for Pedestrian Network Monitoring and Data Extrapolation.In *Transportation Research Board 95th Annual Meeting*, 2016.
[24] Weppner, J., P. Lukowicz, U. Blanke, and G. Troster. Participatory Bluetooth scans serving as urban crowd probes. *Sensors Journal, IEEE,* Vol. 14, No. 12, 2014, pp. 4196-4206.
[25] Weppner, J., and P. Lukowicz. Bluetooth based collaborative crowd density estimation with mobile phones.In *Pervasive computing and communications (PerCom), 2013 IEEE international conference on*, IEEE, 2013. pp. 193-200.
[26] Nicolai, T., and H. Kenn. About the relationship between people and discoverable Bluetooth devices in urban environments.In *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology*, ACM, 2007. pp. 72-78.
[27] Kalogianni, E., R. Sileryte, M. Lam, K. Zhou, M. Van der Ham, S. Van der Spek, and E. Verbree. Passive WiFi monitoring of the rhythm of the campus.In *Proceedings of The 18th AGILE International Conference on Geographic Information Science; Geographics Information Science as an Enabler of Smarter Cities and Communities, Lisboa (Portugal), June 9-14, 2015; Authors version*, Agile, 2015.
[28] Bonne, B., A. Barzan, P. Quax, and W. Lamotte. 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*, IEEE, 2013. pp. 1-6.
[29] Schauer, L., M. Werner, and P. Marcus. 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*, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014. pp. 171-177.
[30] Versichele, M., L. De Groote, M. C. Bouuaert, T. Neutens, I. Moerman, and N. Van de Weghe. Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. *Tourism Management,* Vol. 44, 2014, pp. 67-81.
[31] Danalet, A., B. Farooq, and M. Bierlaire. A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures. *Transportation Research Part C: Emerging Technologies,* Vol. 44, 2014, pp. 146-170.
[32] Vu, L., P. Nguyen, K. Nahrstedt, and B. Richerzhagen. Characterizing and modeling people movement from mobile phone sensing traces. *Pervasive and Mobile Computing,* Vol. 17, 2015, pp.

220-235.
[33] Chilipirea, C., A.-C. Petre, C. Dobre, and M. van Steen. Presumably simple: monitoring crowds using WiFi.
[34] Morgul, E., H. Yang, A. Kurkcu, K. Ozbay, B. Bartin, C. Kamga, and R. Salloum. Virtual Sensors: Web-Based Real-Time Data Collection Methodology for Transportation Operation Performance Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2442, 2014, pp. 106-116.
[35] *Rasberry Pi*. https://www.raspberrypi.org/. Accessed July 11, 2016.
[36] Prasad, S., P. Mahalakshmi, A. Sunder, and R. Swathi. Smart Surveillance Monitoring System Using Raspberry PI and PIR Sensor. *Int. J. Comput. Sci. Inf. Technol,* Vol. 5, 2014, pp. 7107-7109.
[37] Mydlarz, C., J. Salamon, and J. P. Bello. The Implementation of Low-cost Urban Acoustic Monitoring Devices. *arXiv preprint arXiv:1605.08450*, 2016.
[38] Cheng, Y., X. Li, Z. Li, S. Jiang, Y. Li, J. Jia, and X. Jiang. AirCloud: a cloud-based air-quality monitoring system for everyone.In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, ACM, 2014. pp. 251-265.
[39] Abraham, K., and S. Pandian. A Low-Cost Mobile Urban Environmental Monitoring System.In *2013 4th International Conference on Intelligent Systems, Modelling and Simulation*, IEEE, 2013. pp. 659-664.
[40] Leccese, F., M. Cagnetti, and D. Trinca. A smart city application: A fully controlled street lighting isle based on Raspberry-Pi card, a ZigBee sensor network and WiMAX. *Sensors,* Vol. 14, No. 12, 2014, pp. 24408-24424.
[41] Catania, V., and D. Ventura. An approch for monitoring and smart planning of urban solid waste management using smart-M3 platform.In *Open Innovations Association FRUCT, Proceedings of 15th Conference of*, IEEE, 2014. pp. 24-31.
[42] Gaitan, S., L. Calderoni, P. Palmieri, M.-C. ten Veldhuis, D. Maio, and M. B. van Riemsdijk. From sensing to action: Quick and reliable access to information in cities vulnerable to heavy rain. *IEEE Sensors Journal,* Vol. 14, No. 12, 2014, pp. 4175-4184.
[43] Moussaid, M., E. G. Guillot, M. Moreau, J. Fehrenbach, O. Chabiron, S. Lemercier, J. Pettré, C. Appert-Rolland, P. Degond, and G. Theraulaz. Traffic instabilities in self-organized pedestrian crowds. *PLoS Comput Biol,* Vol. 8, No. 3, 2012, p. e1002442.

LIST OF TABLES

LIST OF FIGURES

TABLE 1 Available Pedestrian Counting Technologies (Source: FHWA 2013)

	Technology	Pedestrian Detection	Cost
Permanent	Wi-Fi & Bluetooth Sensors	○	$
↑	Pressure Sensor	○	$$
	Radar Sensor	○	$-$$
How Long?	Seismic sensor	○	$$
	Video Imaging: Automated	○	$-$$
↓	Infrared Sensor (Active or Passive)	●	$-$$
	Video Imaging: Manual	○	$-$$$
Temporary/Short Term	Manual Observers	●	$$-$$$
○	Indicates what is technologically possible		
●	Indicates a common practice		
$, $$, $$	Indicates relative cost per data point		

TABLE 2 Conducted Tests

Study Location	Length	Data	# of Locations
Brooklyn I	5 days	7,792	3
Brooklyn II	1 day	2,159	2
Brooklyn III	5 days	9,755	2
Transit Center	1.5 hours	1,102	2
Commercial Building	14 days	22,759	3
Student Center	1 hour	1,266	3
School of Engineering	2 hours	1,875	4

From/To	**Gate 1**	**Gate 2**	**Gate 3**	**Gate 4**
Entrance A	960	**4487**	4	517
Entrance B	**6444**	5310	81	4

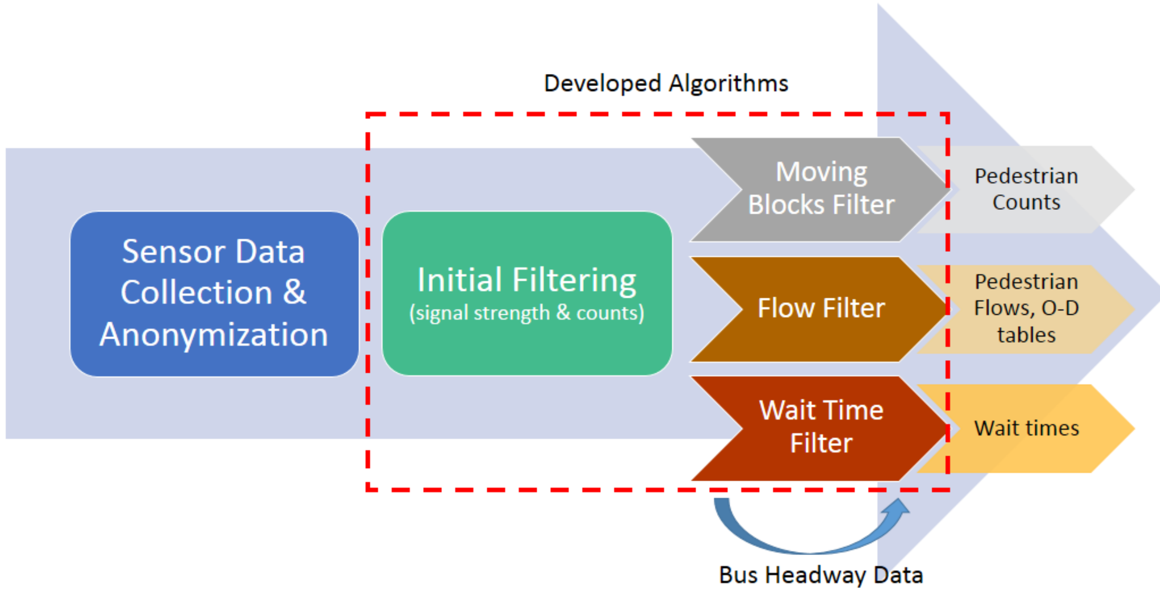


FIGURE 1 Proposed Methodology for Data Analysis

FIGURE 2 Raspberry PI

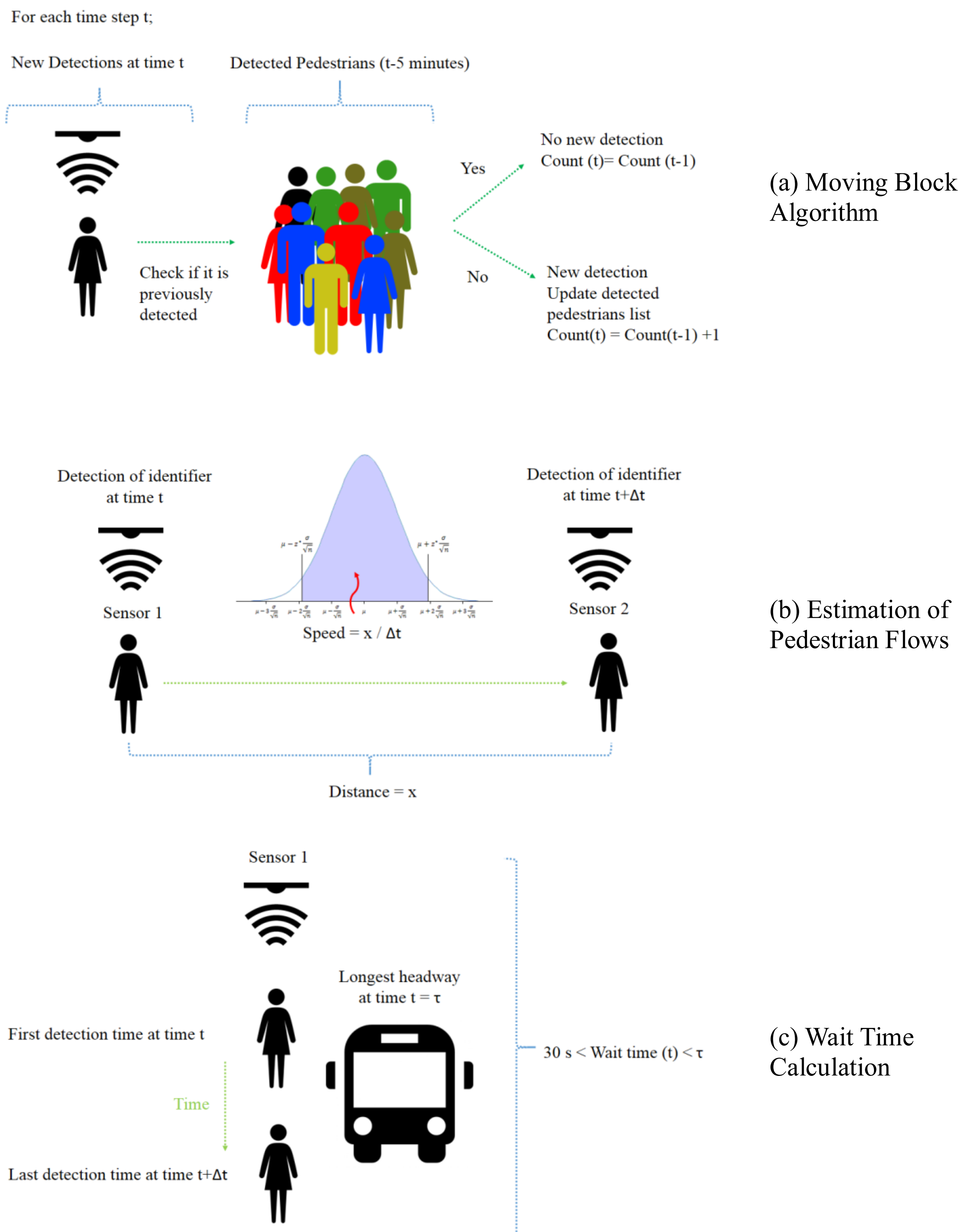
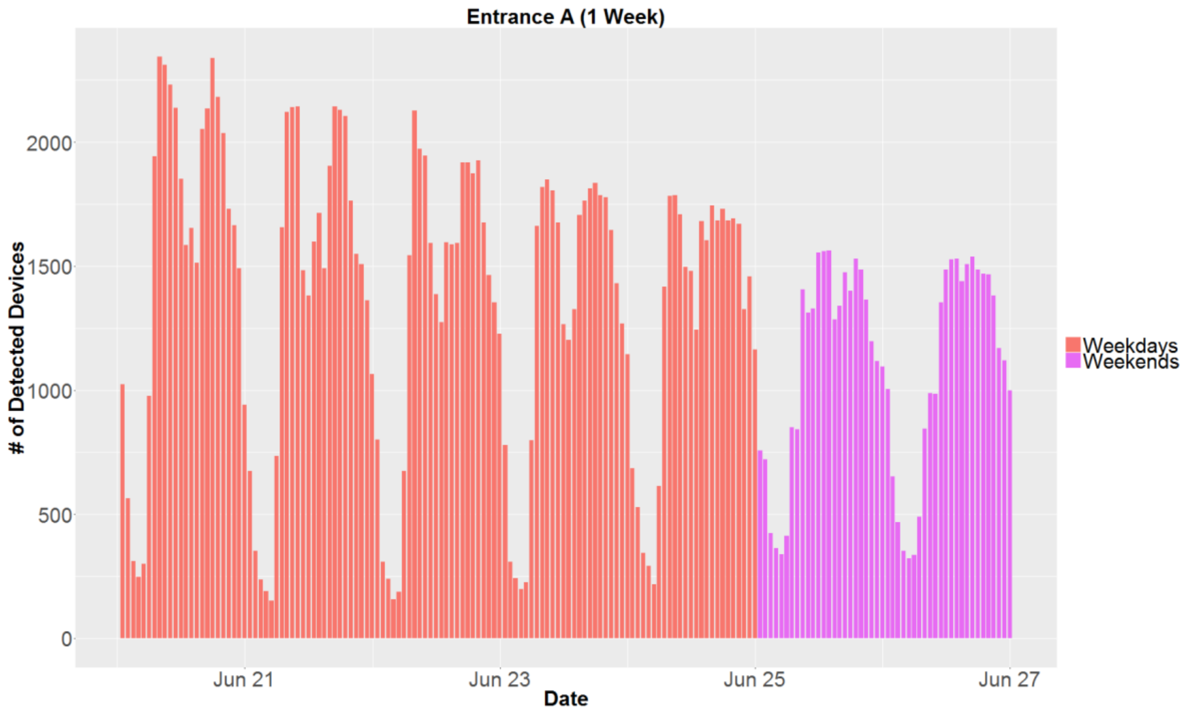
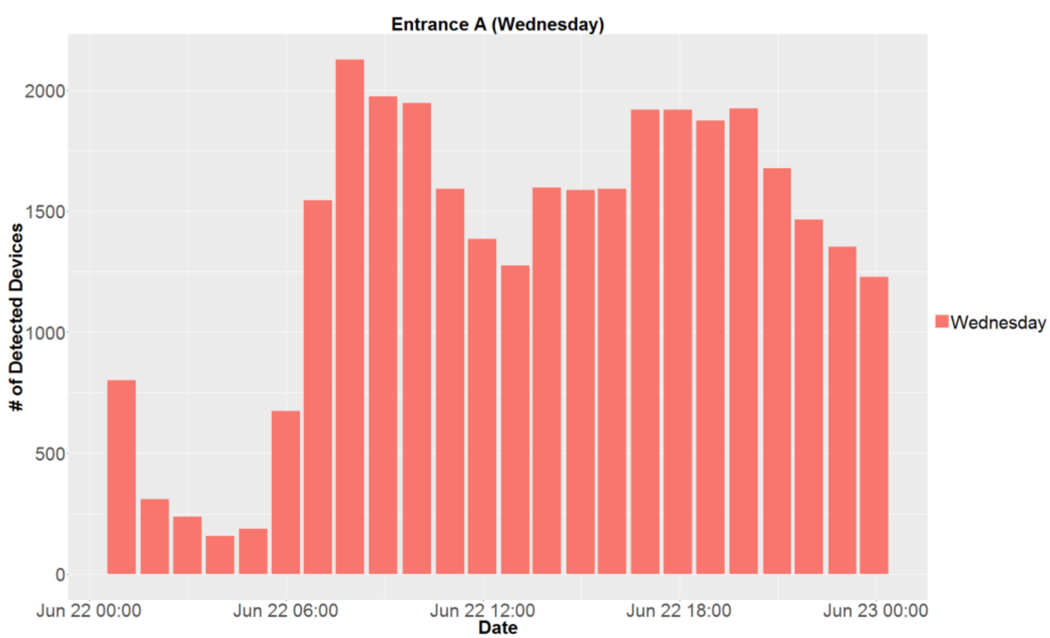


FIGURE 3 Proposed Filtering Algorithms



(a) Detected Mobile Devices for a Week



(b) Detected Devices for a Day

FIGURE 4 Number of Detected Devices

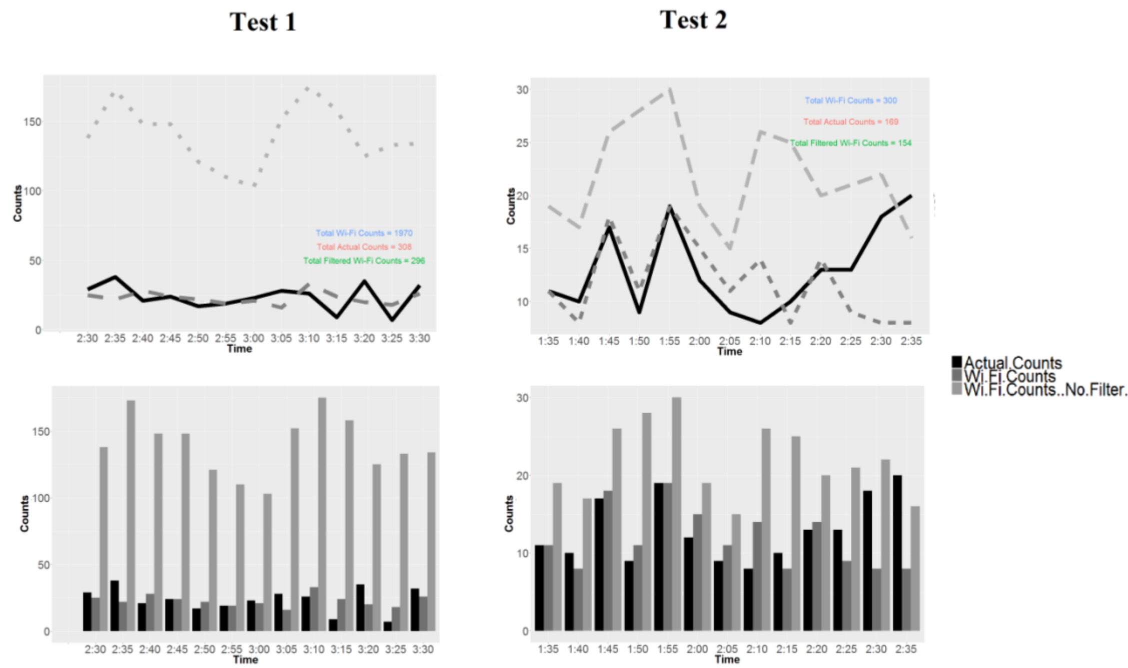


FIGURE 5 Moving Block Algorithm Results

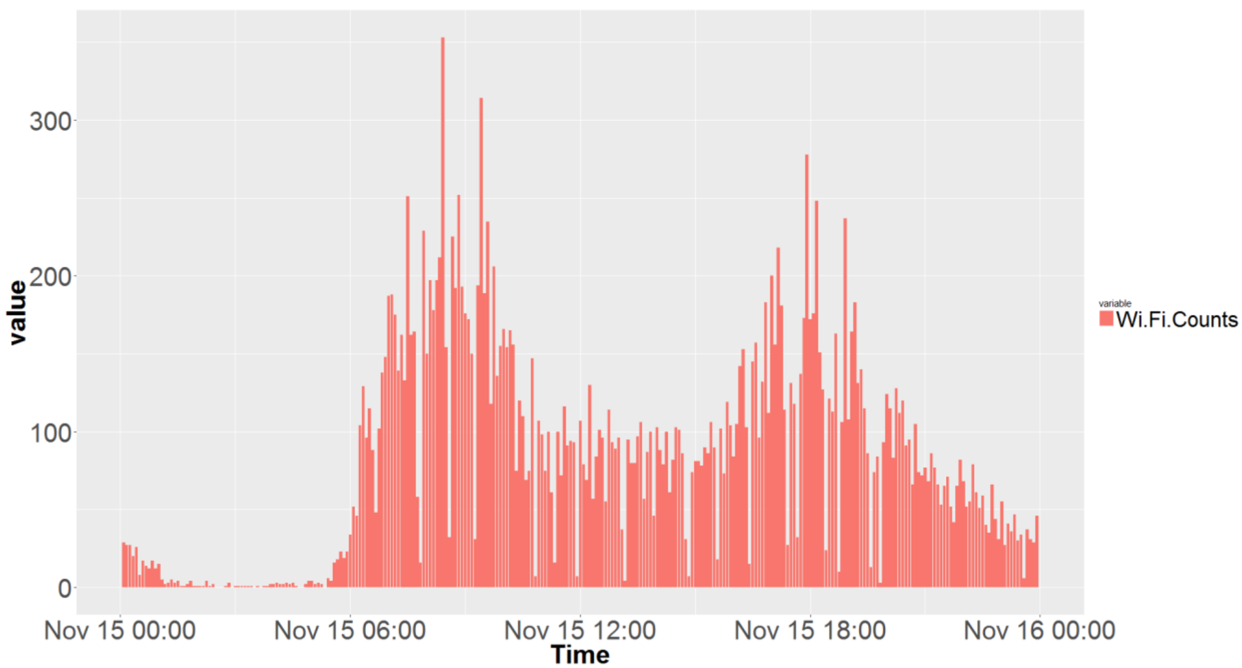


FIGURE 6 Pedestrian Counts for a Day at Entrance A

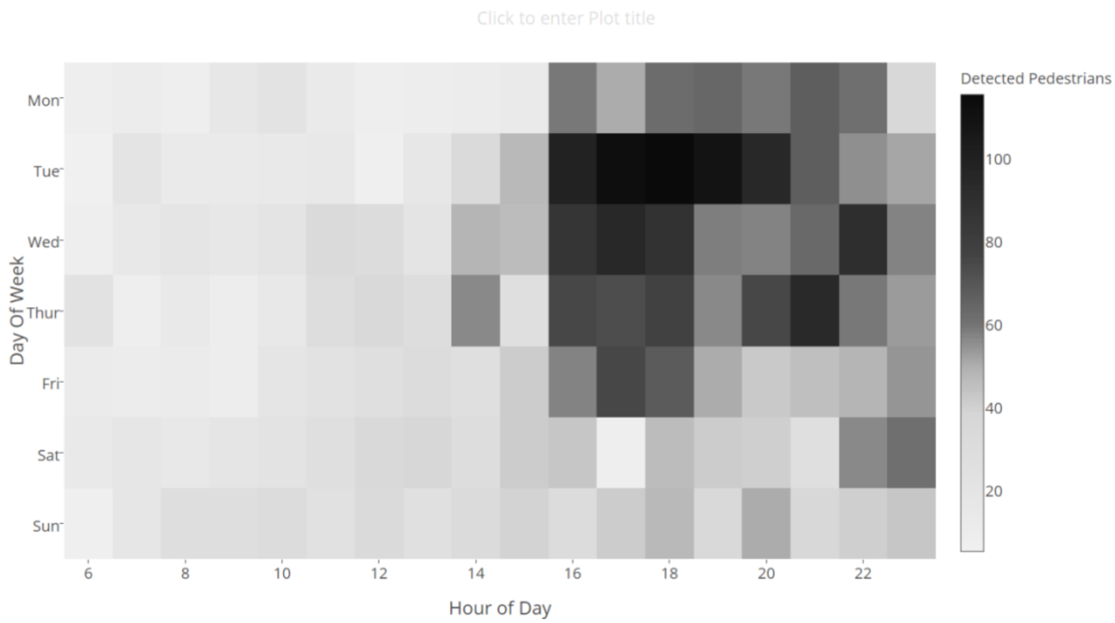


FIGURE 7 Detected Pedestrians between Entrance A and Gate 2

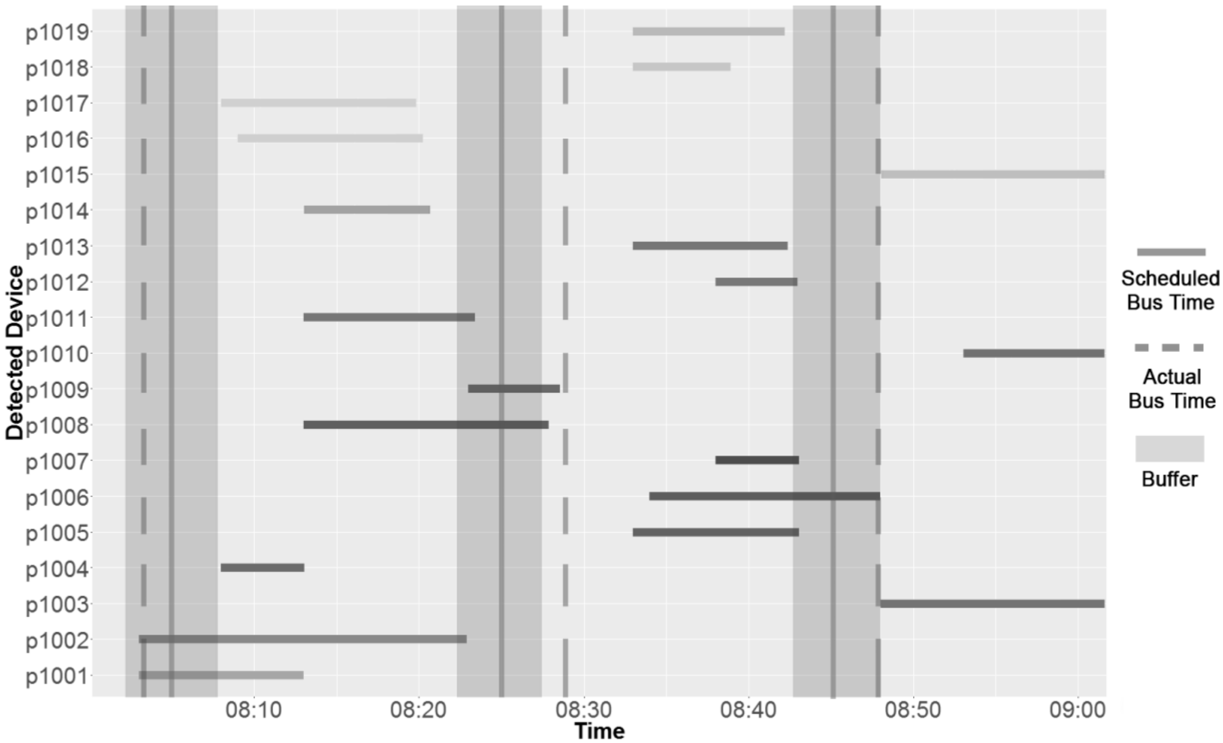


FIGURE 8 Wait Time Visualization