sustainability

MDPI

Article

Evaluation of Bluetooth Detectors in Travel Time Estimation

Krit Jedwanna [1] and Saroch Boonsiripant [2,*]

[1] Department of Civil Engineering, Faculty of Engineering, Rajamangala University of Technology Phra Nakhon, Bangkok 10300, Thailand; krit.j@rmutp.ac.th
[2] Department of Civil Engineering, Faculty of Engineering, Kasetsart University, Bangkok 10900, Thailand
* Correspondence: saroch.b@ku.th; Tel.: +66-81-885-9300

Abstract: With the current popularity of mobile devices with Bluetooth technology, numerous studies have developed methods to analyze the data from such devices to estimate a variety of traffic information, such as travel time, link speed, and origin–destination estimations. However, few studies have comprehensively determined the impact of the penetration rate on the estimated travel time derived from Bluetooth detectors. The objectives of this paper were threefold: (1) to develop a data-processing method to estimate the travel time based on Bluetooth transactional data; (2) to determine the impact of vehicle speeds on Bluetooth detection performance; and (3) to analyze how the Bluetooth penetration rate affected deviations in the estimated travel time. A 28 km toll section in Bangkok, Thailand, was chosen for the study. A number of Bluetooth detectors and microwave radar devices were installed to collect traffic data in October 2020. Five data-processing steps were developed to estimate the travel time. Based on the results, the penetration rate during the day (50 to 90 percent) was higher than during the night (20 to 50 percent). In addition, we found that speed had adverse effects on the MAC address detection capability of the Bluetooth detectors; for speeds greater than 80 km/h, the number of MAC addresses detected decreased. The minimum Bluetooth penetration rate should be at least 1 percent (or 37 vehicles/h) during peak periods and at least 5 percent (or 49 vehicles/h) during the off-peak period.

Keywords: tollway; Bluetooth; detection rate; penetration rate; travel time

check for updates

Citation: Jedwanna, K.; Boonsiripant, S. Evaluation of Bluetooth Detectors in Travel Time Estimation. *Sustainability* **2022**, *14*, 4591. https://doi.org/10.3390/su14084591

Academic Editors: Javier Alonso Ruiz and Angel Llamazares

Received: 27 February 2022
Accepted: 6 April 2022
Published: 12 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Efficient traffic-data collection is necessary to estimate and manage the traffic condition and lessen congested road networks in large developing cities that have toll-road networks and complicated main- and sub-routes. Conventionally, traffic data such as traffic volume, origin–destination data, travel period, and speed of traveling for different time periods are manually collected through surveys. Nowadays, various technologies can be used to collect traffic condition data for most survey research studies. For example, recent studies [1,2] used vehicle trajectory data from smartphone and GPS-equipped vehicles to estimate travel behavior and road traffic conditions. Mobile phone location data are also used to estimate travel time and the origin–destination matrix [3–5]. Among the many choices of detecting devices, with wide ranges in cost and implementation time, low-cost Bluetooth technology is one of the most popular and has been extensively embedded in several essential devices such as smartphones, car stereo speakers, wireless headphones, and other gadgets used in our daily life. Unlike GPS-equipped vehicles, Bluetooth detectors do not rely on individual vehicles to send location data to a central system. Using Bluetooth devices as traffic detectors is also non-intrusive and more affordable compared to other types of sensors. In addition, Bluetooth detectors are less sensitive to low light and inclement weather compared to an automatic license-plate recognition approach [6]. Consequently, Bluetooth detection technology has become a preferred means of major traffic data collection to support the analysis of traffic conditions on many road networks globally.

In this study, conducted on weekdays in October 2020, we evaluated the accuracy of Bluetooth detectors used to collect traffic information on the Don Muang Tollway, Bangkok, Thailand. The penetration rate of Bluetooth detectors was analyzed by comparing the number of trips detected by Bluetooth detectors and the traffic volume counted using microwave radar. In addition, the effect of vehicle speeds on detection capability and the impact of travel time on the accuracy of traffic prediction were investigated.

2. Previous Studies

Bluetooth is a short-range communications technology requiring low power. It has been used to perform wireless communication between fixed and mobile devices within a 10 m range, yet the range of operation may be extended up to 100 m [6,7]. Bluetooth operates at frequencies between 2.402 and 2.480 GHz, or 2.400 and 2.4835 GHz, including guard bands 2 MHz wide at the bottom end and 3.5 MHz wide at the top. The connection is performed by pairing each device through the device's Media Access Control (MAC) address, typically represented as 12 hexadecimal characters (e.g., 60:C4:CC:9D:42:F9). The MAC addresses are assigned by the manufacturers without any publicly available list and centralized database. On the other hand, some hardware MAC addresses are programmable, so it is possible for two devices to have the same MAC address [8,9]. Due to security, Bluetooth Low Energy (BLE) on some devices may periodically use randomized addresses instead of their permanent MAC addresses to prevent tracking on their public channels. Therefore, the same device may announce its presence to other devices with different MAC addresses [10].

A fixed Bluetooth detector will scan the relevant spectrum from the mobile Bluetooth device by broadcasting packets (advertising process). After being detected, the MAC address of the mobile device is recorded and encrypted. Bluetooth transmitters (or detectors) perform transmission (or detection) periodically to find definitive information of the detected device. The raw data received by one detector are limited, but the rest of the data can be collected using more detectors. As the amount of detected data increases significantly, some useful inferences from those data can be made. Traffic data that can be acquired by Bluetooth communication include travel time, travel patterns, travel frequency, stop duration, overtaking behavior, and pedestrian surveys [6,11,12].

On the other hand, telecommunication requirements, privacy issues [13], and low sampling rates are limitations of this technology. One of the most important issues in Bluetooth detection systems is their short detection range. Due to Bluetooth communication installation and other related factors such as signal strength, distance, antennas, and the speed of the vehicle under the detection area, not every Bluetooth device can be successfully detected. In [14], it was revealed that only 80 percent of the Bluetooth devices in vehicles can be detected correctly while passing the detecting zone.

One study [15] highlighted that the penetration rate of the Bluetooth detector is an important factor that defines the quality of the estimation. It concluded that the penetration rate should be at least 3 percent of the total vehicle volume and the minimum time-period used to collect traffic data that can provide reliable travel-time estimation is 20 days. Another study [16] proposed a detection probability model for moving Bluetooth devices. Empirical measurement of the Bluetooth penetration rate has been reported in a few studies [17,18], and [19] and the average penetration rate values reported were 4.5 percent, 29 percent, and 4.3–6.9 percent, respectively.

Two previous studies [20,21] collected the Bluetooth detection rates for the highways in Europe and the USA. They found that the penetration rate largely fluctuated in the range of 0.2–70 percent. The average detection rate reported in Bavaria, Germany, was 25 percent, and in Europe it was 30 percent. The Bluetooth matching rate is generally in the 5-percent range for most areas in the USA. In the Netherlands, the detection rates for the freeways A6, A7, A31, and A32 were rather high (54–70 percent). The detection rates from different places depend on various factors such as the type of sensors, antennas, position and angle of the installation, road configuration, and traffic characteristics.

From [19], accurate traffic information requires the determination of a minimum appropriate sample size. Although some studies concluded that 5 percent was a minimum acceptable sampling rate, other factors, such as the level of accuracy and roadway conditions, must be considered while determining the minimum sampling rate. Thus, the minimum sample size for an entire set of roadway links is not constant and can be obtained dynamically. The sample-size requirements may also depend on the time of day. For exceptionally low or very high traffic volumes, the minimum percentage sample rate should be higher than that for medium levels of traffic volume.

The work in [19] studied the vehicle detection rate for Bluetooth traffic sensors in Maryland and Delaware in the USA. It was found that peak/off-peak hours and traffic volume significantly affected the penetration rate. Bluetooth detectors can detect vehicles more successfully during peak hours due to vehicles moving more slowly due to the high traffic volume. One study [20] found that any Bluetooth detector could detect vehicles traveling at a speed below 120 km/h. The Maryland study [19] also found that the Bluetooth penetration rate during the period from midnight to 7:00 a.m. was higher than during other periods, which coincided with the high freight transportation activity during the night. The authors in [19] concluded that unfiltered data from a single vehicle might contain multiple devices. Thus, the detection rate result could be higher than usual. A similar study [20] explored the A3 Freeway in Germany and reported that the detection rate among trucks was in the range of 65–70 percent, indicating that they were detected approximately five times more often than passenger cars that had a detection rate in the range of 12–17 percent.

In addition to the Bluetooth penetration rate, data processing and filtering are equally important [22]. On the Don Muang Tollway, it is likely that a single transit vehicle has several devices, especially for public transit vehicles, such as intercity public vans and buses. The Bluetooth detector would detect more than one MAC addresses from the same vehicle, which might create bias in the sampling process. Therefore, a data-cleaning and -filtering process is necessary to ensure the accuracy of the estimated travel time from such technology.

3. Data Collection

3.1. Study Corridor

The Don Muang Tollway is a 28 km toll road connecting central Bangkok with its northern suburbs and serving approximately 120,000 vehicles/day. The toll-road structure is elevated 14 m above the at-grade street network in the area.

3.2. Bluetooth Detector Setup

Along its length, the Don Muang Tollway has 11 stations installed with two Bluetooth detectors/station for inbound and outbound traffic. The average distance between adjacent stations is in the range 0.5–4 km, as shown in Figure 1. The specification and technology of the Bluetooth detectors are given in Table 1.

Table 1. Specifications of Bluetooth detectors.

Item	Descriptions
Detector	Bluetooth 4.0 Class 1
Maximum coverage range	100 m
Sampling rate	1 Hz
Maximum speed detection	120 km/h
Storage	32 GB
Antenna type	Flat panel 9 dB
Frequency range	2.402~2.480 GHz
Transmit Output Power	+19 dBm (+6 dBm EDR) E.I.R.P.
Antenna Connector	RP-SMA
Operating temperature	−40~+85 °C
Power consumption	10 W
Interface	Ethernet 100 mbps
Power supply	Power over Ethernet 802.11af

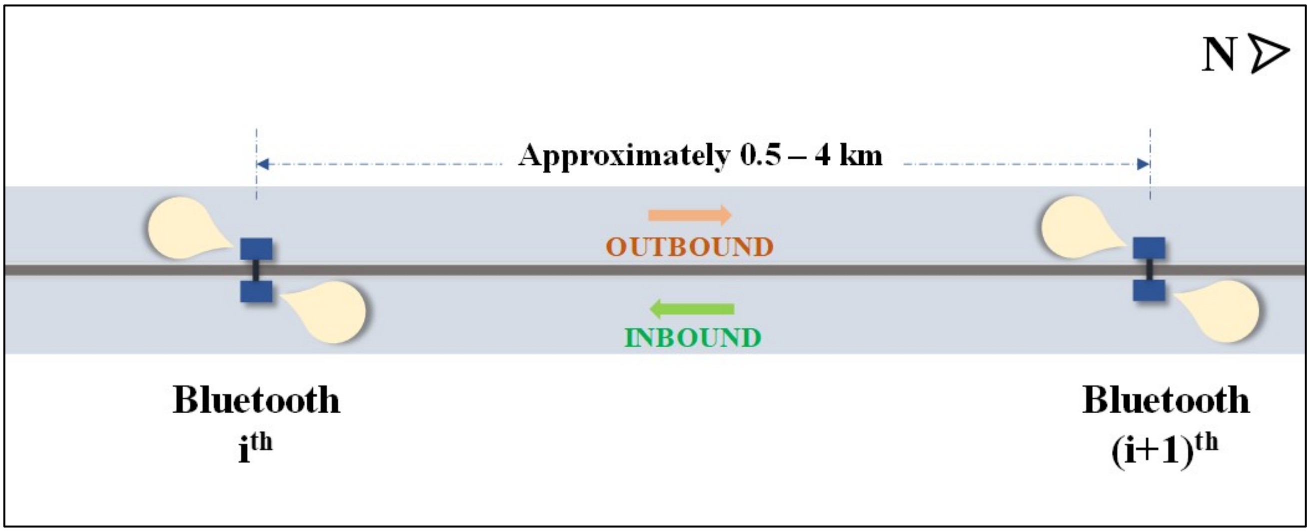

Figure 1. Bluetooth detector installation configuration.

3.3. Location/Dataset

To determine the reference for the detection, fourteen microwave radar detectors were installed along the 28 km road corridor to count passing traffic as a baseline for Bluetooth penetration rate determination. The locations of all Bluetooth detector stations and microwave sensors on the Don Muang Tollway are illustrated in Figure 2.

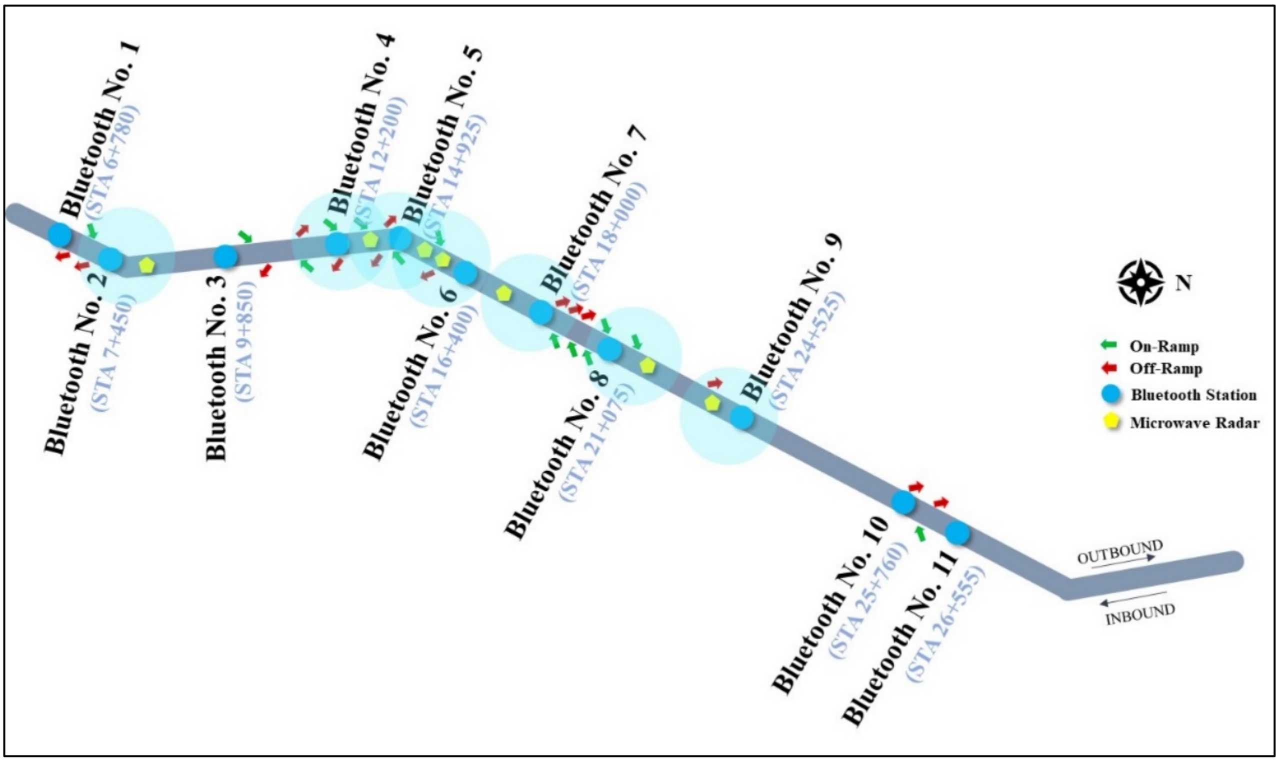

Figure 2. Locations of Bluetooth detectors (blue filled circles) and microwave radars (yellow filled pentagons) along study corridor.

3.4. Bluetooth Detector Data

MAC addresses and timestamp data from the Bluetooth devices were collected at each of the 11 Bluetooth stations for 20 weekdays during October 2020. Data obtained from each Bluetooth detector consisted of Record ID, Device ID, Timestamp, MAC address, and other related information as described in Table 2. Notably, the MAC address provides a unique identifier for each Bluetooth device. In addition, the Timestamp was recorded once the MAC address had been successfully detected.

Table 2. Data fields obtained from Bluetooth detectors.

Variable	Description
Record ID	A number in the sequence of detection
Device ID	Identification number of the detector
Timestamp	The time whenever the device MAC Address is detected
MAC address	MAC address of the Bluetooth device
First timestamp	The first time when the device MAC address is detected
Last timestamp	The last time when the device MAC address is detected

An example of the dataset is shown in Table 3**.** In this work, comprehensive data-analysis was performed on 19.4 million records collected from all stations for 20 weekdays.

Table 3. Sample results from devices passing Bluetooth detectors.

Record ID	Device ID	Timestamp	MAC Address
114779044	Bluetooth No. 9	00:00.0	5A:72:74:06:94:DB
114779045	Bluetooth No. 9	00:00.0	59:EE:9C:71:37:56
114779046	Bluetooth No. 9	00:00.1	77:5B:F3:89:4A:34
114779047	Bluetooth No. 9	00:00.1	7A:A9:8D:A7:0E:5E
114779048	Bluetooth No. 9	00:00.1	58:49:75:FD:A5:2D
114779049	Bluetooth No. 9	00:00.1	54:D8:1E:51:6B:29
114779050	Bluetooth No. 9	00:00.2	5A:8D:A8:D9:AE:85
114779051	Bluetooth No. 9	00:02.9	41:F4:97:27:9F:E6
114779052	Bluetooth No. 9	00:03.7	6D:B4:D3:D9:8D:29
114779053	Bluetooth No. 9	00:04.1	7B:28:B1:51:A0:9F
114779054	Bluetooth No. 9	00:04.2	67:F0:36:FB:8E:23
114779055	Bluetooth No. 9	00:04.8	4E:47:EF:8D:03:7A
114779056	Bluetooth No. 9	00:04.9	EB:BC:50:FC:A3:5B

Later, the Device ID, Timestamp, and MAC address data were used to analyze the travel time, number of detected MAC addresses, and the Bluetooth penetration rate. For example, travel time can be determined as the difference between Timestamps of the same MAC address between the two Device IDs on the study corridor. Furthermore, the Bluetooth penetration rate can be calculated as the ratio of the number of detected MAC addresses to the number of vehicles obtained from the nearby microwave radar. Data-cleaning and -filtering processes were also required, and these are discussed in the next section.

3.5. Traffic Volume Data Collection

The reference traffic volume in this study was identified by the installed microwave radar devices as shown in Table 4. Each of collected set of data consisted of *Device ID*, *Reftime*, *Updatetime*, *Lane*, and *Totalcount*. The *Lane* of the toll road was identified by one of the numbers one (slowest), two, or three (fastest), with lane number zero denoting all three travel lanes. *Totalcount* is the number of vehicles captured every minute. Notably, the microwave radar units used in this study had been recently installed and calibrated. The volume count accuracy was at least 95 percent.

Table 4. Sample traffic volume data collected using microwave radar.

Device ID	Updatetime	Lane	Totalcount
TCU-BK-RDO-02	2020-09-30T17:00:25.550Z	3	0
TCU-BK-RDO-02	2020-09-30T17:00:25.551Z	0	2
TCU-BK-RDO-02	2020-09-30T17:00:25.551Z	2	1
TCU-BK-RDO-02	2020-09-30T17:00:25.551Z	1	1
TCU-BK-RDO-02	2020-09-30T17:01:25.792Z	3	2
TCU-BK-RDO-02	2020-09-30T17:01:25.792Z	0	4
TCU-BK-RDO-02	2020-09-30T17:01:25.792Z	2	2
TCU-BK-RDO-02	2020-09-30T17:01:25.793Z	1	0
TCU-BK-RDO-02	2020-09-30T17:02:26.134Z	3	2
TCU-BK-RDO-02	2020-09-30T17:02:26.135Z	0	3
TCU-BK-RDO-02	2020-09-30T17:02:26.135Z	2	1
TCU-BK-RDO-02	2020-09-30T17:02:26.135Z	1	0

According to the traffic counts from the microwave radar devices in October 2020, the inbound and outbound traffic volumes had different patterns, with the inbound traffic peaking from 6:00 a.m. to 9:00 a.m. and the outbound direction characterized by a peak from 4:00 p.m. to 7:00 p.m., as shown in Figure 3. This pattern indicates the characteristics of commuter traffic between the northern suburbs and central Bangkok served by the Don Muang Tollway.

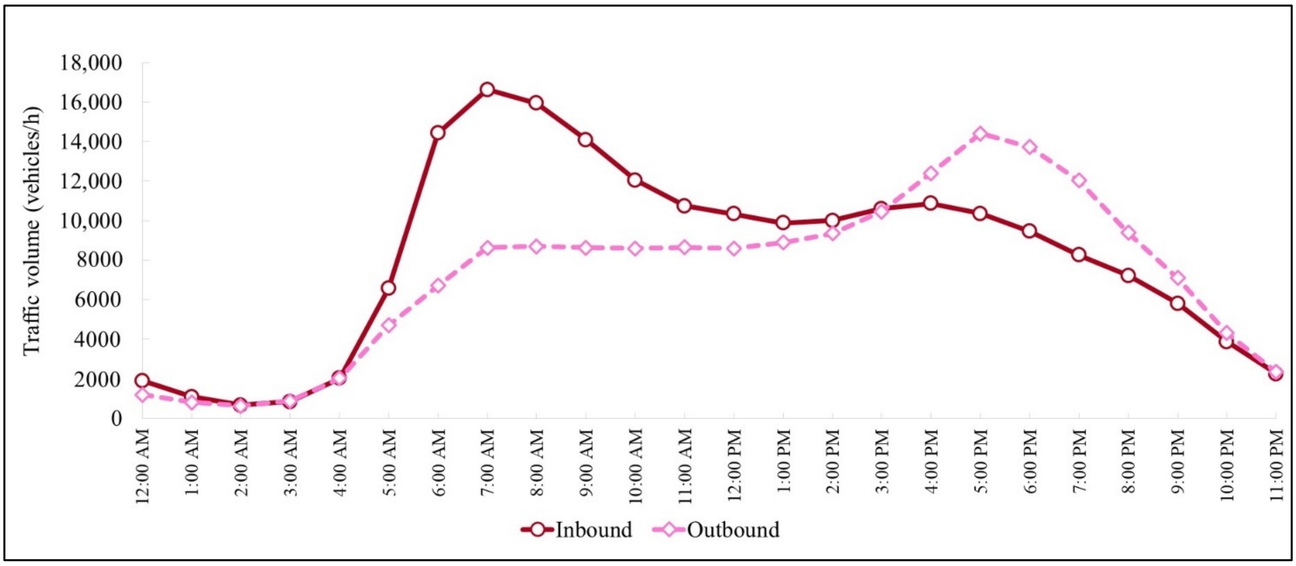

Figure 3. Hourly traffic volume collected from microwave radar device for inbound and outbound traffic directions for weekdays in October 2020.

4. Data Processing and Analysis

4.1. Data Processing

As mentioned in [20], one MAC address does not necessarily represent an individual vehicle because there might be multiple Bluetooth devices in a vehicle. Furthermore, driving behavior and other factors can cause errors in data collection that affect the penetration-rate analysis. Therefore, data-cleaning and -filtering processes were required before performing further analysis. In the data processing, duplicate MAC addresses were removed to obtain only one value representing a vehicle, which was then used to perform node-to-node (station-to-station) analysis, and consequently to create a trip model for each MAC address.

Figure 4 shows the five important steps of data processing of raw data collected from the Bluetooth detectors. The main objective of data processing was to achieve trip data for one vehicle trip that would be suitable for the penetration rate analysis. The PostgreSQL database management system was used for data cleaning based on the following steps:

1. Duplicate data cleaning;
2. MAC addresses matching between Bluetooth detectors;
3. Direction identification;
4. Trip determination;
5. Same-vehicle-trip removal.

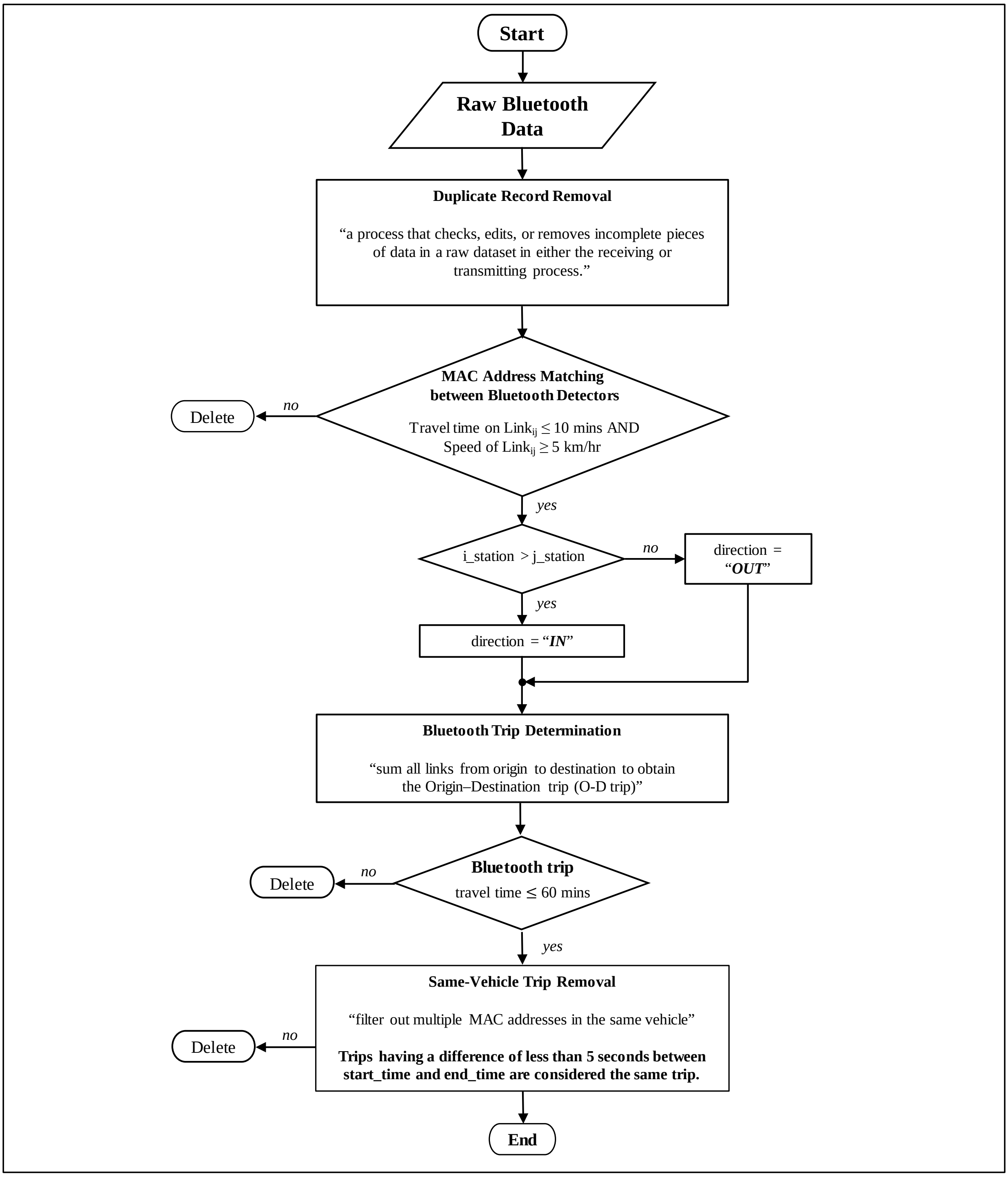

Figure 4. Data cleaning and filtering processing.

4.1.1. Data Cleaning

The raw transactional data obtained from the 22 Bluetooth detectors, including *device_id*, *bt_station*, *ref_time*, *date_bt*, *time_bt*, and *mac_address*, are described in Table 5. From the initial dataset of 19.4 million records, we removed approximately 18 percent as duplicate records. Additionally, the MAC addresses detected at only a single Bluetooth location could not be used to estimate the travel time on the toll road; thus, such MAC addresses were removed, resulting in 16 million records remaining for the next step.

Table 5. Data field descriptions in raw data collection.

Variable	Description
id	A number in the sequence of detection
device_id	Identification number of the detector and direction
bt_station	The name of the detector
ref_time	The date and time whenever the device MAC Address is detected
date_bt	The date whenever the device MAC Address is detected
time_bt	The time whenever the device MAC Address is detected
mac_address	MAC address of the Bluetooth device

4.1.2. MAC Address Matching between Bluetooth Detectors

Each record in the raw dataset refers to a data point where the MAC addresses were detected at one Bluetooth station. In this step, we connected these data points and converted them into link data.

First, the data were sorted by the three attributes *mac_address*, *ref_time*, and *device_id*, respectively. Next, we defined $link_{ij}$ for each MAC address as a pair of records with the same MAC address between stations *i* and *j*, as shown in Figure 5.

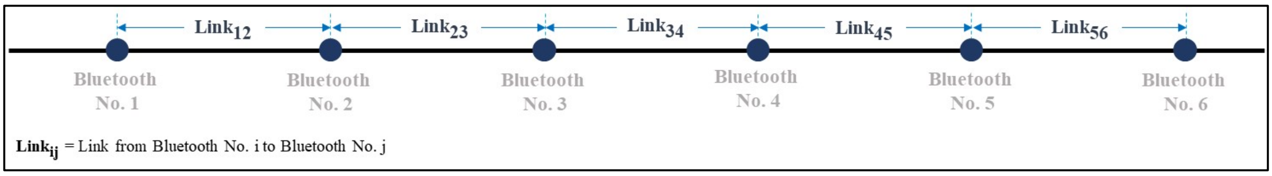

Figure 5. Definition of $link_{ij}$.

To create $link_{ij}$ data, the record (k + 1) was copied and appended to the existing record *k*. Figure 6 represents the processed data in this step. The data table is composed of three parts “node *i*” data, “node *j*” data, and “$link_{ij}$” data. “Node *i*” data consist of *mac_address*, *i_date*, *i_time*, and *i_station*, where “node *j*” data consist of *mac_address*, *j_date*, *j_time*, and *j_station*. Lastly, the $link_{ij}$ column contains a pair of Bluetooth station numbers at which the MAC address was detected. Data attributes are described in Table 6. It is possible that the same vehicle creates multiple trips on the tollway. Therefore, the time-difference threshold has been specified to determine separate trips with the same MAC address. The time difference between the *i*th and *j*th stations was calculated. If the time difference was more than 10 min, the link was discarded. Notably, a time difference of 10 min equals the 99th percentile of all travel times between adjacent Bluetooth stations, which is equivalent to a travel speed of 5 km/h or less.

Figure 6 demonstrates how each $link_{ij}$ was constructed for the MAC address 1C:91:9D:F2:5E:32. At line 1, this MAC address was initially detected by Bluetooth No. 10 at time 7:00:00 prior to Bluetooth No. 9 at time 7:01:05. Thus, the link between the two stations was defined as $link_{10\text{-}9}$. Similarly, in line 2, the same MAC address was detected by Bluetooth No. 9 (*i_station* = 9) and Bluetooth No. 8 (*j_station* = 8). As a result, $link_{9\text{-}8}$ was constructed to represent the MAC address movement between the two stations.

Node i					Node j				Link$_{ij}$
Line No.	mac_address	i_date	i_time	i_station	mac_address	j_date	j_time	j_station	link_ij
1	1C:91:9D:F2:5E:32	2020-10-01	7:00:00	10	1C:91:9D:F2:5E:32	2020-10-01	7:01:05	9	10-09
2	1C:91:9D:F2:5E:32	2020-10-01	7:01:05	9	1C:91:9D:F2:5E:32	2020-10-01	7:04:16	8	09-08
3	1C:91:9D:F2:5E:32	2020-10-01	7:05:00	8	1C:91:9D:F2:5E:32	2020-10-01	7:08:23	7	08-07
4	1C:91:9D:F2:5E:32	2020-10-01	7:08:24	7	1C:91:9D:F2:5E:32	2020-10-01	7:10:02	6	07-06
5	1C:91:9D:F2:5E:32	2020-10-01	7:10:02	6	1C:91:9D:F2:5E:32	2020-10-01	7:11:24	5	06-05
6	1C:91:9D:F2:5E:32	2020-10-01	7:11:24	5	1C:91:9D:F2:5E:32	2020-10-01	7:13:45	4	05-04
7	1C:91:9D:F2:5E:32	2020-10-01	7:13:46	4	1C:91:9D:F2:5E:32	2020-10-01	7:15:49	3	04-03
8	1C:91:9D:F2:5E:32	2020-10-01	7:15:49	3	1C:91:9D:F2:5E:32	2020-10-01	7:17:56	2	03-02
9	1C:91:9D:F2:5E:32	2020-10-01	7:17:56	2	1C:91:9D:F2:5E:32	2020-10-01	7:18:40	1	02-01
10	1C:91:9D:F2:5E:32	2020-10-01	13:00:51	1	1C:91:9D:F2:5E:32	2020-10-01	13:01:27	2	01-02
11	1C:91:9D:F2:5E:32	2020-10-01	13:01:28	2	1C:91:9D:F2:5E:32	2020-10-01	13:03:21	3	02-03
12	1C:91:9D:F2:5E:32	2020-10-01	13:03:21	3	1C:91:9D:F2:5E:32	2020-10-01	13:05:08	4	03-04
13	1C:91:9D:F2:5E:32	2020-10-01	13:05:09	4	1C:91:9D:F2:5E:32	2020-10-01	13:07:11	5	04-05
14	1C:91:9D:F2:5E:32	2020-10-01	13:07:11	5	1C:91:9D:F2:5E:32	2020-10-01	13:08:20	6	05-06
15	1C:91:9D:F2:5E:32	2020-10-01	13:08:20	6	1C:91:9D:F2:5E:32	2020-10-01	13:09:31	7	06-07
16	1C:91:9D:F2:5E:32	2020-10-01	13:09:32	7	1C:91:9D:F2:5E:32	2020-10-01	13:11:56	8	07-08
17	1C:91:9D:F2:5E:32	2020-10-01	13:12:00	8	1C:91:9D:F2:5E:32	2020-10-01	13:14:37	9	08-09
18	1C:91:9D:F2:5E:32	2020-10-01	13:14:37	9	1C:91:9D:F2:5E:32	2020-10-01	13:15:34	10	09-10
19	1C:91:9D:F2:5E:32	2020-10-01	13:15:37	10	1C:91:9D:F2:5E:32	2020-10-01	13:16:53	11	10-11

Figure 6. Example of $link_{ij}$ table for MAC address matching between two Bluetooth stations.

Table 6. Data attributes used in determining $link_{ij}$.

Variable	Description
mac_address	MAC address of the detected Bluetooth device
i_date	The date when the MAC address is detected at Bluetooth No. i
i_time	The timestamp when the MAC Address is detected at Bluetooth No. i
i_station	The name of detector detected at Bluetooth No. i
j_time	The timestamp when the MAC Address is detected at Bluetooth No. j
j_date	The date when the MAC Address is detected at Bluetooth No. j
j_station	The name of detector at Bluetooth No. j
$link_{ij}$	Link between Bluetooth numbers i and j with the same MAC address

4.1.3. Direction Identification

The traffic characteristics for different traffic directions vary. Therefore, it is important to identify the MAC address direction before determining travel-time values from this dataset. To identify the MAC address direction of each $link_{ij}$, we began by sorting the link data using date/time information (*i_date*, *i_time*, *j_date*, and *j_time*, respectively). Then, the values of the *i_station* and *j_station* were compared. If *i_station* is greater than *j_station*, the direction is inbound (southbound). On the other hand, if *i_station* is less than *j_station*, this indicates that the MAC Address direction is outbound (northbound). Figure 7 shows direction identification for $link_{ij}$. Notably, the *i_date/i_time* is always prior to the *j_date/j_time* since the records were sorted according to date/time information in Section 4.1.2.

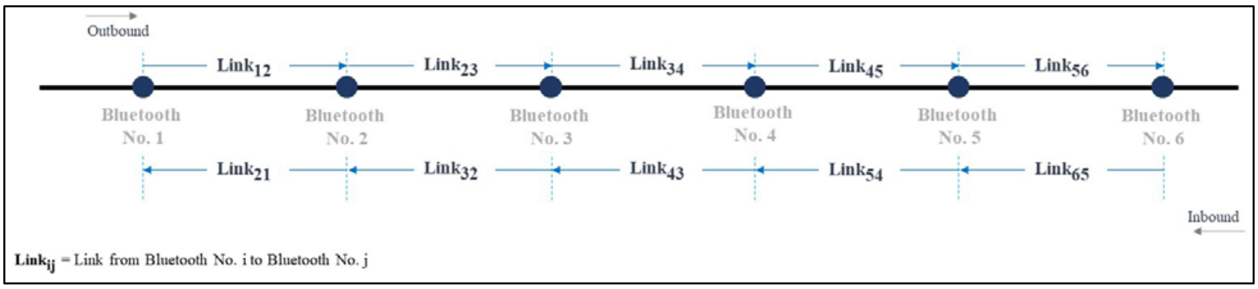

Figure 7. Direction of $link_{ij}$.

Figure 8 gives an example of how to obtain the traffic direction for a vehicle with MAC address 1C:91:9D:F2:5E:32. This MAC address was detected by Bluetooth No. 10 (*i_station* = 10) at time 7:00:00 and then detected again by Bluetooth No. 9 (*j_station* = 9) at the time 7:01:05. Therefore, the MAC address direction of $\text{link}_{10\text{-}9}$ is from Bluetooth No. 10 to No. 9, which can be identified as the inbound direction.

Node i | Node j | Link_{ij} | i_station > j_station, direction = IN

Line No.	mac_address	i_date	i_time	i_station	mac_address	j_date	j_time	j_station	link_ij	direction
1	1C:91:9D:F2:5E:32	2020-10-01	7:00:00	10	1C:91:9D:F2:5E:32	2020-10-01	7:01:05	9	10-09	IN
2	1C:91:9D:F2:5E:32	2020-10-01	7:01:05	9	1C:91:9D:F2:5E:32	2020-10-01	7:04:16	8	09-08	IN
3	1C:91:9D:F2:5E:32	2020-10-01	7:05:00	8	1C:91:9D:F2:5E:32	2020-10-01	7:08:23	7	08-07	IN
4	1C:91:9D:F2:5E:32	2020-10-01	7:08:24	7	1C:91:9D:F2:5E:32	2020-10-01	7:10:02	6	07-06	IN
5	1C:91:9D:F2:5E:32	2020-10-01	7:10:02	6	1C:91:9D:F2:5E:32	2020-10-01	7:11:24	5	06-05	IN
6	1C:91:9D:F2:5E:32	2020-10-01	7:11:24	5	1C:91:9D:F2:5E:32	2020-10-01	7:13:45	4	05-04	IN
7	1C:91:9D:F2:5E:32	2020-10-01	7:13:46	4	1C:91:9D:F2:5E:32	2020-10-01	7:15:49	3	04-03	IN
8	1C:91:9D:F2:5E:32	2020-10-01	7:15:49	3	1C:91:9D:F2:5E:32	2020-10-01	7:17:56	2	03-02	IN
9	1C:91:9D:F2:5E:32	2020-10-01	7:17:56	2	1C:91:9D:F2:5E:32	2020-10-01	7:18:40	1	02-01	IN
10	1C:91:9D:F2:5E:32	2020-10-01	13:00:51	1	1C:91:9D:F2:5E:32	2020-10-01	13:01:27	2	01-02	OUT
11	1C:91:9D:F2:5E:32	2020-10-01	13:01:28	2	1C:91:9D:F2:5E:32	2020-10-01	13:03:21	3	02-03	OUT
12	1C:91:9D:F2:5E:32	2020-10-01	13:03:21	3	1C:91:9D:F2:5E:32	2020-10-01	13:05:08	4	03-04	OUT
13	1C:91:9D:F2:5E:32	2020-10-01	13:05:09	4	1C:91:9D:F2:5E:32	2020-10-01	13:07:11	5	04-05	OUT
14	1C:91:9D:F2:5E:32	2020-10-01	13:07:11	5	1C:91:9D:F2:5E:32	2020-10-01	13:08:20	6	05-06	OUT
15	1C:91:9D:F2:5E:32	2020-10-01	13:08:20	6	1C:91:9D:F2:5E:32	2020-10-01	13:09:31	7	06-07	OUT
16	1C:91:9D:F2:5E:32	2020-10-01	13:09:32	7	1C:91:9D:F2:5E:32	2020-10-01	13:11:56	8	07-08	OUT
17	1C:91:9D:F2:5E:32	2020-10-01	13:12:00	8	1C:91:9D:F2:5E:32	2020-10-01	13:14:37	9	08-09	OUT
18	1C:91:9D:F2:5E:32	2020-10-01	13:14:37	9	1C:91:9D:F2:5E:32	2020-10-01	13:15:34	10	09-10	OUT
19	1C:91:9D:F2:5E:32	2020-10-01	13:15:37	10	1C:91:9D:F2:5E:32	2020-10-01	13:16:53	11	10-11	OUT

j_station > i_station, direction = OUT

Figure 8. Example of direction identification for $link_{ij}$.

4.1.4. Trip Determination

In this step, we constructed each trip by summing all links from the origin to the destination to obtain the origin–destination trip (O-D trip); that is, a trip made up of a collection of the links determined in Section 4.1.2. Figure 9 illustrates the trip derived from the sequence of links.

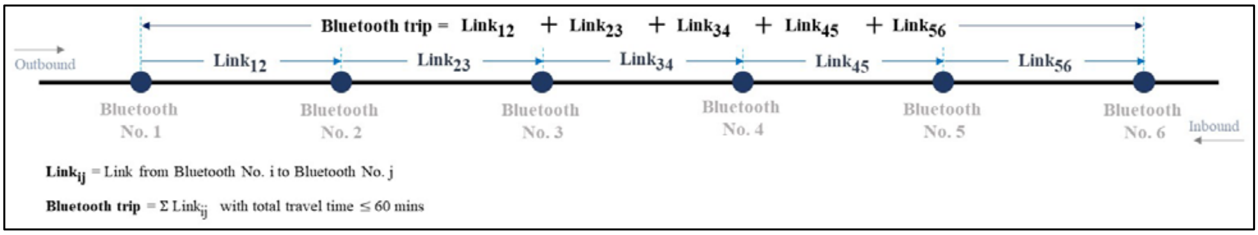

Figure 9. Single trip defined as series of consecutive links.

Data for a single trip consist of mac_address, trip_path (a list of Bluetooth numbers), direction, start_time (timestamp at the first Bluetooth detector), end_time (timestamp at the last Bluetooth detector), and dif_time (time difference in minutes between start_time and end_time). An example of trip data is shown in Figure 10. Notably, the total travel time is checked and if it exceeds 60 min the algorithm removes the trip, as this implies an average speed of less than 28 km/h, which is unlikely to happen on the 28 km tollway.

	mac_address	trip_path	direction	start_time	end_time	dif_time
Inbound Trip	1C:91:9D:F2:5E:32	10-09-08-07-06-05-04-03-02-01	IN	2020-10-01 7:00:00	2020-10-01 7:18:40	18.67
Outbound Trip	1C:91:9D:F2:5E:32	01-02-03-04-05-06-07-08-09-10-11	OUT	2020-10-01 13:00:51	2020-10-01 13:16:53	16.03

Figure 10. Example of trip determination results.

For example, the MAC address shown in Figure 8 (lines 1–9) can be used to construct a trip with the following information:

1. *mac_address* = 1C:91:9D:F2:5E:32;
2. *trip_path* = 10-09-08-07-06-05-04-03-02-01;
3. *direction*: IN;
4. *start_time* = 2020-10-01 7:00:00;
5. *end_time* = 2020-10-01 7:18:40;
6. *dif_time* = 18.67 min.

As a result, this vehicle spent 18.67 min traveling inbound from Bluetooth No. 10 at 07:00 a.m. to Bluetooth No. 1 at 07:18 a.m.

4.1.5. Same-Vehicle-Trip Removal

It is possible that one vehicle contains more than one MAC address, especially for high-occupancy vehicles such as transit vans and buses. This could create biases in the dataset; for example, it could result in high-occupancy vehicles having a greater influence in the travel-time dataset. Therefore, it was important to detect multiple MAC addresses potentially in the same vehicle and apply a proper treatment to obtain a single sample as the representative of the individual vehicle.

Logically, if multiple MAC addresses are detected at a Bluetooth detector at nearly the same time (less than a 5 s time window) and the same group of MAC addresses are detected at the next Bluetooth detector also at nearly the same time, it is likely that these MAC addresses are from the same vehicle.

Therefore, we developed an algorithm to detect if multiple trips (with different MAC addresses) had differences in both *start_time* and *end_time* of less than 5 s; if so, they were considered as a single trip to represent the movement of one vehicle [20].

To identify multiple trips that were potentially from the same vehicle, we reordered the trips according to time, grouped these trips into the criterion of a 5 s difference, and indexed the representative of all the other trips with the *group_id* (the red line in Figure 11). The *start_time* (first timestamp) of the representative is the minimum value and the *end_time* (last timestamp) is the maximum value of all trips in the group. Figure 11 on the left shows that trips from Bluetooth No.1 to No. 2 and No. 3 of MAC addresses A (Path: A_1-A_2-A_3), B (Path: B_1-B_2-B_3), and C (Path: C_1-C_2-C_3) are the same. On the other hand, Figure 11 on the right illustrates a complete trip that stems from only one Bluetooth device/vehicle.

4.2. Penetration Rate Calculation

In our work, the penetration rate (PR) at a Bluetooth station is defined as the ratio of the number of trips detected by the Bluetooth detector and the traffic volume confirmed by the nearby microwave radar devices, as given in Equation (1):

$$\text{Penetration rate (PR)} = \frac{\text{No. of trips passing at a Bluetooth station}}{\text{No. of vehicle detected by microwave radar}} \tag{1}$$

To calculate Bluetooth penetration rates along the study corridor, the numbers of detected MAC addresses yielded from the data processing steps were compared with traffic counts from the nearby microwave sensor stations where there were no on/off ramps in between the pairs of Bluetooth detectors and microwave radar devices. Based on the availability of the microwave radar devices, Bluetooth numbers 5, 7, 8, and 9 were selected for the inbound traffic direction and Bluetooth numbers 4, 6, and 8 were selected for the outbound traffic direction, as summarized in Table 7.

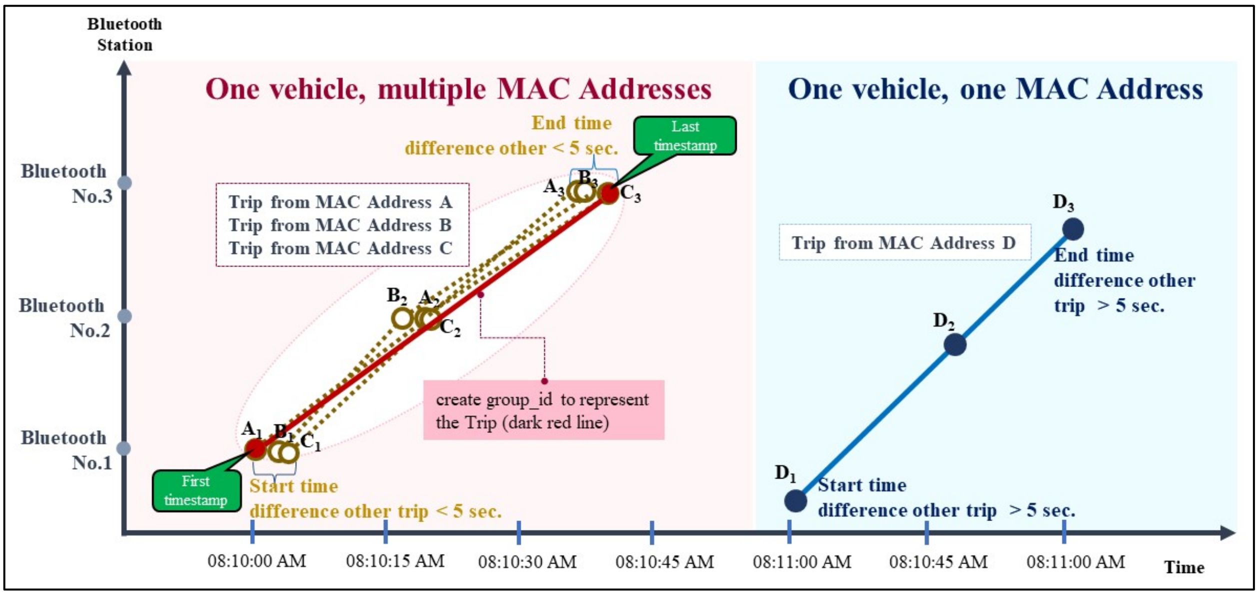

Figure 11. Example of same-vehicle-trip removal.

Table 7. Bluetooth stations with nearby microwave radar for penetration-rate analysis.

Direction	Bluetooth		Microwave Radar		Distance (km)
	Station ID	Location	Station ID	Location	
Inbound	Bluetooth No.5	14 + 952	TCU-BK-RDO-01	15 + 280	0.3
	Bluetooth No.7	18 + 000	TCU-BK-RDO-03	17 + 370	0.6
	Bluetooth No.8	21 + 075	TCU-DM-RDS-02	21 + 451	0.4
	Bluetooth No.9	24 + 525	TCU-AN-RDO-01	23 + 300	1.2
Outbound	Bluetooth No.4	12 + 200	TCU-LP-RDO-01	12 + 525	0.3
	Bluetooth No.6	16 + 400	TCU-BK-RDO-02	15 + 750	0.6
	Bluetooth No.8	21 + 075	TCU-DM-RDS-01	21 + 450	0.4

4.3. Travel-Time Statistics Calculation

In this study, two statistics were calculated every 15 min to represent travel time used on the toll road, namely the 50th percentile travel time and the 85th percentile travel time. The former represents the median value of the travel-time distribution, while the latter also captures travel-time dispersion [23]. The 85th percentile travel time is currently chosen to display on the variable message signs in the facility which represents the travel time that 85 percent of the drivers spend to complete their journey on the toll road.

5. Results and Discussion

5.1. Data Processing Results

Based on the proposed data-cleaning and -filtering process, Table 8 illustrates the amount of data remaining after each step. Based on the 20 weekdays of data collection, there were 19.4 million records resulting from 11 Bluetooth detector stations. Approximately 40 percent of the total records remained from the data-cleaning and -filtering process.

Table 8. Summary of data-processing results.

No.	Step	Results
1	Raw Data	19.4 million records
2	Duplicate Records Removal	16 million unique records (82% remaining)
3	MAC Address Matching between Bluetooth Detectors	8.7 million links created
4	Direction Identification	4.3 million northbound links 4.4 million southbound links
5	Trip Determination	1.2 million northbound trips 1.2 million southbound trips
6	Same-Vehicle-Trip Removal	0.94 million northbound trips (78%) 1.06 million southbound trips (88%)

During the 20-day period, there were 2 million Bluetooth trips resulting from the data-processing effort. Figure 12 represents the average number of Bluetooth trips/h during weekdays in October 2020. From the figure, the peak time for the inbound direction was from 6:00 AM to 9:00 AM, while the peak time for the outbound direction was from 4:00 PM to 7:00 PM. The resulting Bluetooth trip data were compared with the traffic counts recorded from nearby microwave sensors to calculate Bluetooth penetration rates in the next section.

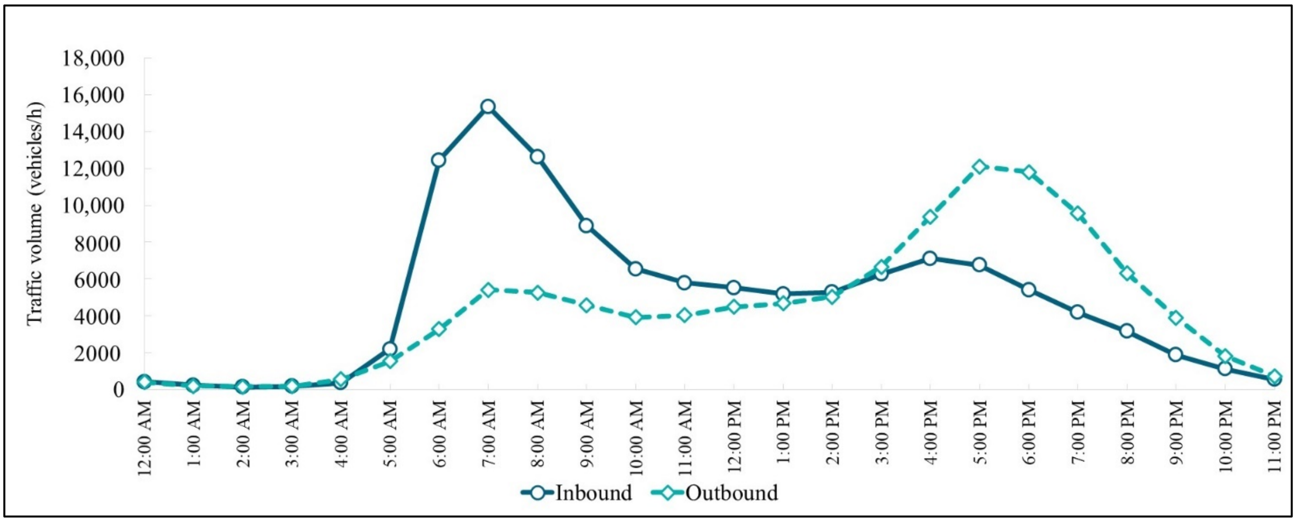

Figure 12. Numbers of Bluetooth trips/h for inbound and outbound traffic directions for weekdays in October 2020.

5.2. Penetration Rate Results

The results of the penetration rate analysis are shown in Figure 13 (inbound traffic) and Figure 14 (outbound traffic). Notably, the penetration rates in each direction are the average hourly penetration rates of the Bluetooth stations with nearby microwave radar devices (Bluetooth numbers 5, 7, 8, and 9 for the inbound traffic and Bluetooth numbers 4, 6, and 8 for the outbound traffic). Each line in these figures represents the Bluetooth penetration rate for a specific day of the week. For inbound traffic, the penetration rates on the study corridor were in the range of 50–90 percent during the day and 20–50 percent during the night. The penetration rates during the morning peak (6:00 a.m.–9:00 a.m.) and evening peak (4:00 p.m.–7:00 p.m.) were also higher than for the off-peak period.

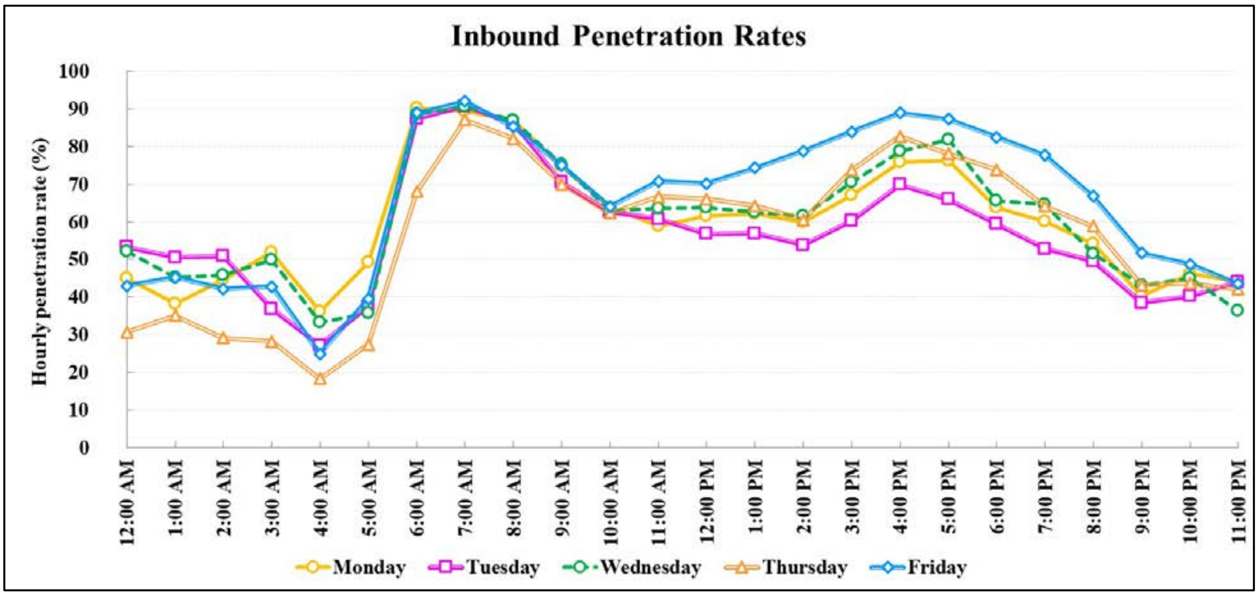

Figure 13. Hourly penetration rates during weekdays; averaged from Bluetooth numbers 5, 7, and 8 for inbound direction.

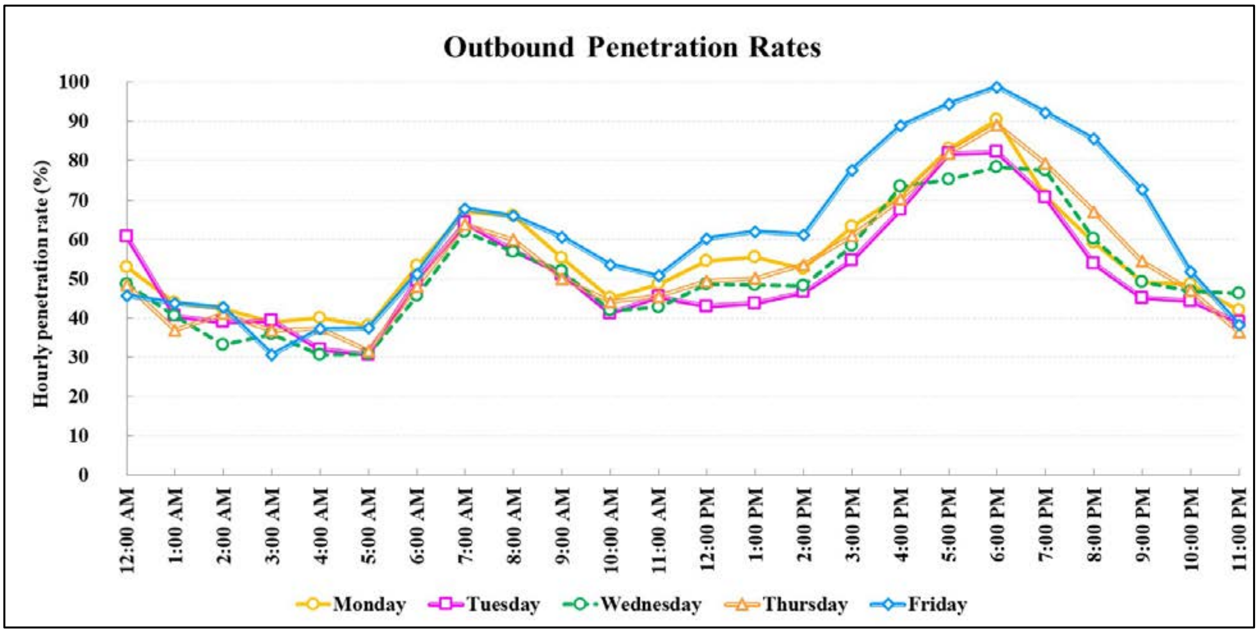

Figure 14. Hourly penetration rates during weekdays; averaged from Bluetooth numbers 5, 7, and 8 for outbound direction.

Outbound traffic also had a similar pattern; the penetration rates peaked during the morning (6:00 a.m.–9:00 a.m.) and evening (4:00 p.m.–7:00 p.m.). The penetration rates during the peak periods ranged from 50 to nearly 100 percent.

The penetration rates were highest during the peak periods—morning peak for the inbound direction (see in Figure 13) and evening peak for the outbound direction (see in Figure 14). Hypothetically, slower traffic during peak periods makes it more likely for the Bluetooth detectors to detect MAC addresses. In Section 5.3, we investigate the relationship between speed and the number of MAC addresses recorded by the Bluetooth detectors.

5.3. Impact of Speed on Bluetooth Detection

From the previous section, the number of MAC addresses and Bluetooth penetration rates were higher during peak periods when traffic flow was high and average speed was low. The free-flow speed measured using microwave radar near Bluetooth No. 8 was 80 km/h. However, during peak periods, the average speed reduced to approximately 60 km/h (a 25 percent reduction). To determine the impact of traffic speeds on Bluetooth detection capability, the 3.1 km section from Bluetooth No. 9 to Bluetooth No. 8 was selected, since it contained no on/off-ramps. Therefore, the number of detected MAC addresses should be similar at both Bluetooth stations. We also used nearby microwave sensors to measure the speed and flow at these two stations, with microwave radar No. 9 located 1.3 km downstream of Bluetooth No.9 and microwave radar No.8 located 0.3 km upstream of Bluetooth No. 8, as shown in Figure 15.

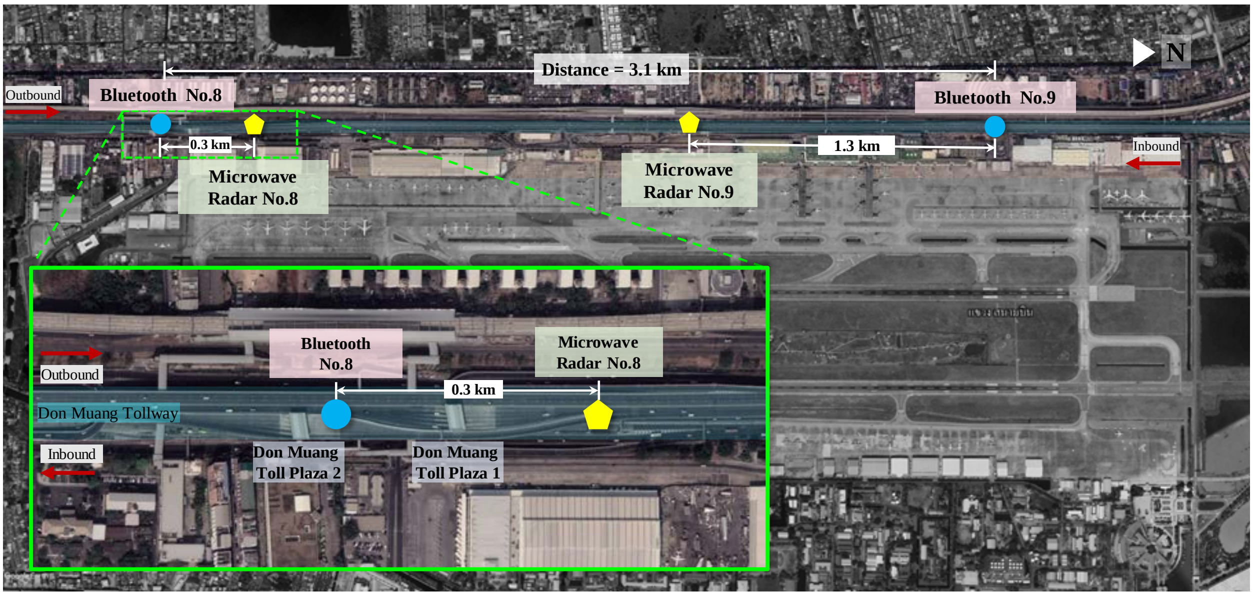

Figure 15. Locations of inbound Bluetooth numbers 8 and 9, microwave radars devices, and Don Muang toll plazas. Traffic moves from Bluetooth number 9 (upstream) to Bluetooth number 8 (downstream).

Inbound traffic data on Monday 26 October 2020 were selected for this analysis because there is typically high inbound traffic volume on Monday mornings. Figure 16 depicts 5 min average speeds at microwave radar numbers 9 (green line) and 8 (red line) throughout the day. Speeds measured at microwave radar number 9 (green line) were relatively stable (90–100 km/h) throughout the day. On the other hand, speeds measured at microwave radar number 8 (red line) ranged between 70 and 80 km/h with a speed drop (60 km/h) during the morning peak (6:00 a.m. to 7:30 a.m.). Speed at the downstream location (Bluetooth number 8) was lower than at the upstream location because the downstream location had two staggered toll plazas where drivers were required to slow down to enter toll booths.

In addition to the speed data, the numbers of MAC addresses detected at Bluetooth numbers 9 (upstream) and 8 (downstream) were analyzed. Since there is no on-ramp or off-ramp between the two stations, the numbers of MAC addresses at both stations should be similar. However, the number of MAC addresses detected at the downstream location was consistently higher than for the upstream location throughout the day. The difference of number of MAC addresses between Bluetooth numbers 8 and 9 (N_{BT8}-N_{BT9}) is illustrated in Figure 16 (gray line). During the night, the difference was nearly zero. However, when the average speed dropped approximately 25 percent at the downstream location during the morning peak, the number of MAC addresses at Bluetooth No. 8

also increased, implying that Bluetooth detection performance increased as the vehicle speed decreased.

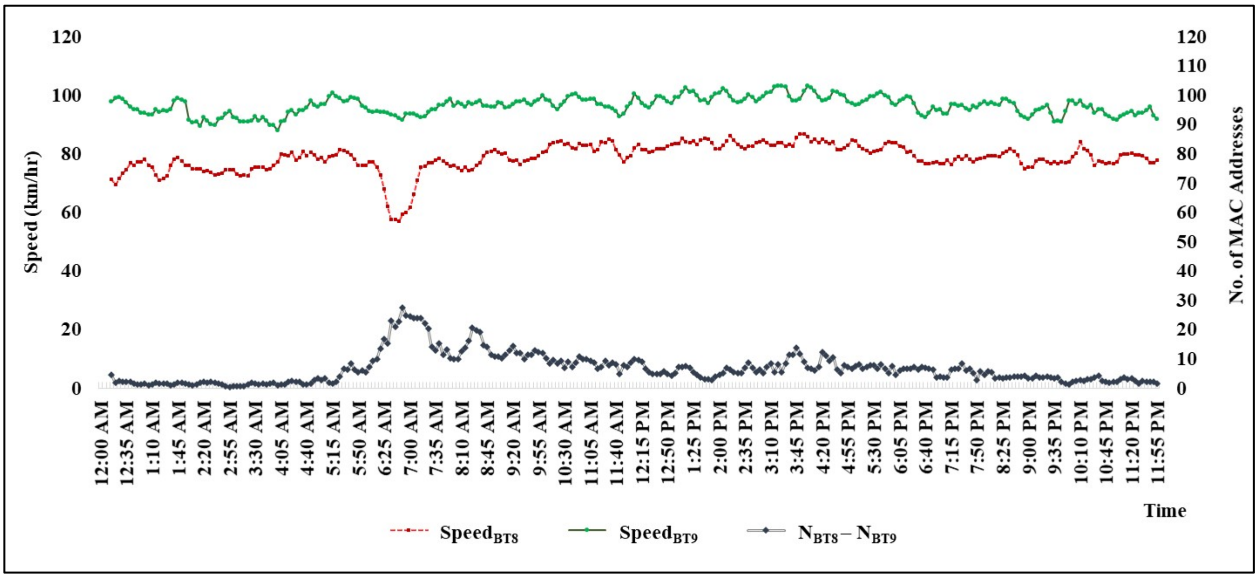

Figure 16. Five-minute average speeds at upstream (green line) and downstream (red line) locations, along with differences in number of MAC addresses (BT8-BT9).

5.4. Sensitivity Analysis of Bluetooth Penetration Rate

We analyzed how various levels of Bluetooth penetration rates affected travel-time estimation. The mean absolute percentage error (MAPE) is commonly used to measure the accuracy of estimated travel times from Bluetooth detectors compared with ground-truth travel times. However, in this study, the actual (ground-truth) travel time was not available. We assumed that the higher penetration rate had a higher accuracy due to its larger sample size. Therefore, we compared travel-time values at different penetration rates, with the travel time at the highest penetration rate for each scenario. Changes in travel-time values affected by various penetration rates were compared with the travel time derived from the highest penetration rate for that scenario. Thus, we modified the traditional MAPE and renamed it the mean absolute percentage deviation (*MAPD*) using the formula:

$$MAPD = \frac{1}{n}\sum\left|\frac{\acute{t}_i - \hat{t}_i}{\acute{t}_i}\right| \times 100\% \tag{2}$$

where n is number of observations, $\acute{t}_i$ is an estimated value from the highest penetration rate data, and $\hat{t}_i$ is an estimated value from lower penetration data.

It is important to select a road segment that contains no on-ramp or off-ramp between a microwave radar and a Bluetooth detector in order to accurately control the number of vehicles in the penetration-rate calculation. For the inbound direction, Bluetooth trips from Bluetooth numbers 9 to 7 were selected. For the outbound direction, Bluetooth trips from Bluetooth numbers 2 to 4 were selected and any trips that passed both the origin and destination points were considered. For example, a trip passing Bluetooth numbers 2, 3, and 4 as well as a trip passing only Bluetooth numbers 2 and 4 would be included in the sensitivity analysis.

It was noted that the penetration rates calculated in this section were generally lower than the penetration rates shown in Section 5.2 because the number of Bluetooth trips in Section 5.2 included all trips that began, passed, or ended at a specific Bluetooth station, while the Bluetooth trips in this section included only trips that passed both the origin and destination in a specific order (from Bluetooth numbers 9 to 7 and from Bluetooth numbers 2 to 4).

Table 9 shows the penetration rates considered for the inbound traffic dataset and Table 10 shows the penetration rates considered for the outbound traffic dataset. Traffic data on Monday 26 October 2020 were used for the inbound analysis and traffic data on Friday 2 October 2020 were considered for the outbound analysis. The analysis contained three time-periods (AM peak, PM peak, and off-peak). Each period had different penetration rates from 1 percent to the highest penetration rate of the dataset. Since the highest penetration rates for each scenario/period were different, the number of penetration rate levels was also different.

Table 9. Inbound scenarios for sensitivity analysis of penetration rates, from Bluetooth numbers 9 to 7 on Monday 26 October 2020.

Period	Actual Penetration Rate (%)	Traffic Volume from Microwave Radar (Vehicles/Period)	Penetration Rate (%)
AM peak (6:00 a.m.–9:00 a.m.)	50	7966	1, 2, 3, 5, 10, 20, 30, 40
PM peak (4:00 p.m.–7:00 p.m.)	41	5618	1, 2, 3, 5, 10, 20, 30, 40
Off peak	32	12,327	1, 2, 3, 5, 10, 20, 30

Table 10. Outbound scenarios for sensitivity analysis of penetration, from Bluetooth numbers 2 to 4 on Friday 2 October 2020.

Period	Actual Penetration Rate (%)	Traffic Volume from Microwave Radar (Vehicles/Period)	Penetration Rate (%)
AM peak (6:00 a.m.–9:00 a.m.)	28	5013	1, 2, 3, 5, 10, 20
PM peak (4:00 p.m.–7:00 p.m.)	29	10,318	1, 2, 3, 5, 10, 20
Off peak	28	17,757	1, 2, 3, 5, 10, 20

For each period, Monte Carlo sampling was used to construct a dataset with a specific penetration rate. That is, Bluetooth trips were randomly drawn with replacement from the total Bluetooth trip dataset until the desired penetration rate was reached for each period. Total traffic volumes and the specified numbers of Bluetooth trips are illustrated in Table 11 (for inbound traffic) and Table 12 (for outbound traffic). The sampling procedure was repeated 10 times for each penetration rate. The 50th and 85th percentile travel-time values were calculated from the sampling trips to represent travel time in each period and penetration levels. The percentile values from the 10 runs were averaged to obtain a representative travel time for the scenario. The MAPD was calculated to determine the difference between the travel time obtained from the highest penetration rate and the travel time obtained from the simulated penetration rate.

Table 11. Number of Bluetooth trips required for sensitivity analysis of inbound trips from Bluetooth numbers 9 to 7 on Monday 26 October 2020.

Period	Traffic Volume (Vehicles/Period)	Number of Bluetooth Trips at Each Penetration Level							
		1%	2%	3%	5%	10%	20%	30%	40%
AM peak (6:00 a.m.–9:00 a.m.)	7966	80	159	239	398	797	1593	2390	3186
PM peak (4:00 p.m.–7:00 p.m.)	5618	56	112	169	281	562	1124	1685	2247
Off peak	12,327	123	247	370	616	1233	2465	3698	

Table 12. Number of Bluetooth trips required for sensitivity analysis of outbound trips from Bluetooth numbers 2 to 4 on Friday 2 October 2020.

Period	Traffic Volume (Vehicles/Period)	Number of Bluetooth Trips at Each Penetration Level							
		1%	2%	3%	1%	10%	20%	1%	40%
AM peak (6:00 a.m.–9:00 a.m.)	5013	50	100	150	251	501	1003		
PM peak (4:00 p.m.–7:00 p.m.)	10,318	103	206	310	516	1032	2064		
Off peak	17,757	178	355	533	888	1776	3551		

The next two sections discuss the sensitivity of the two travel-time statistics, namely the 50th percentile and the 85th percentile travel times, to various penetration rates. The 50th percentile travel time represents the middle value of road users' travel times, while the 85th percentile travel time represents the duration for which 85 percent of the road users could traverse the section.

5.5. Impact of Penetration Rate on Estimated 50th Percentile Travel Time

The results of the sensitivity analysis for the 50th percentile travel times are shown in Figure 17 (inbound traffic) and Figure 18 (outbound traffic).

Notably, we considered an MAPD value of 15 percent as the acceptable threshold currently used by the toll road agency. For the inbound direction, the AM peak had the highest traffic volume (inbound commuting traffic to the city center in the morning) and its MAPD was under 15 percent, even with the lowest penetration rate of 1 percent (with a sample size of 26 MAC addresses/h). During the PM peak (with relatively less traffic) the MAPD exceeded the 15 percent threshold only when the penetration rate was below 1.5 percent (equivalent to 37 MAC addresses/h). In contrast, off-peak required at least 5 percent of the Bluetooth penetration rate to maintain the MAPD value of 15 percent (equivalent to 34 MAC addresses/h).

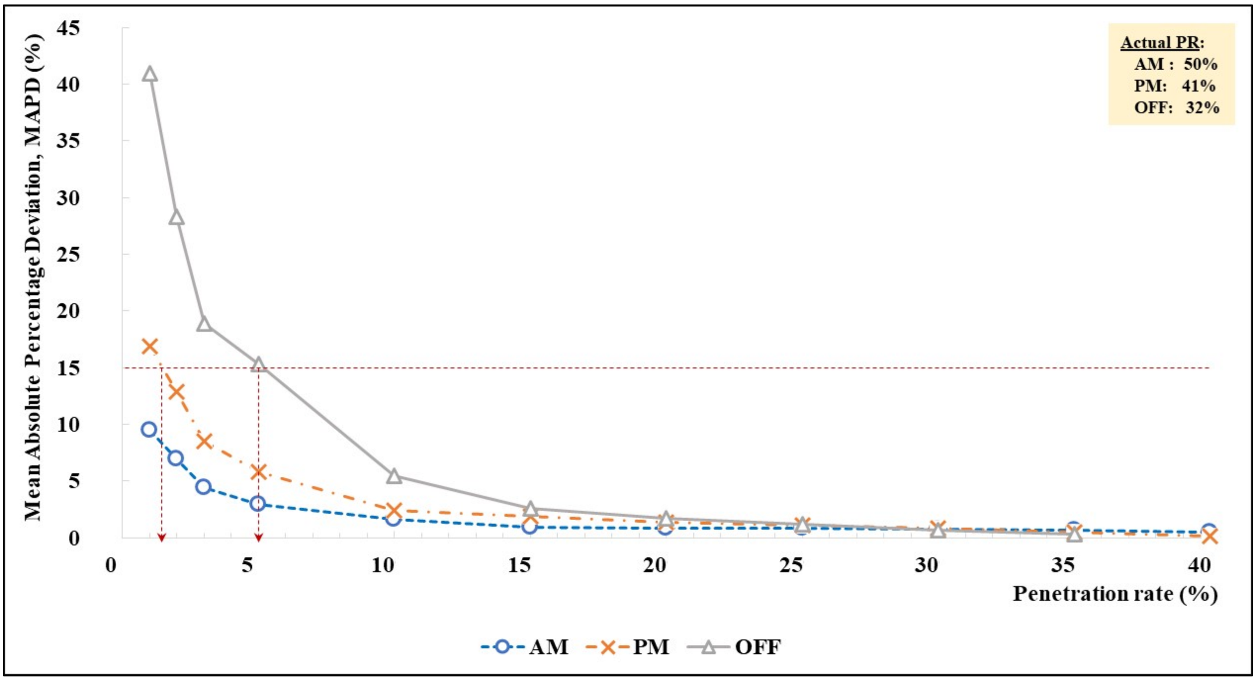

Figure 17. Mean Absolute Percentage Deviation (MAPD) of 50th percentile travel time at different penetration rates for inbound direction on Monday 26 October 2020.

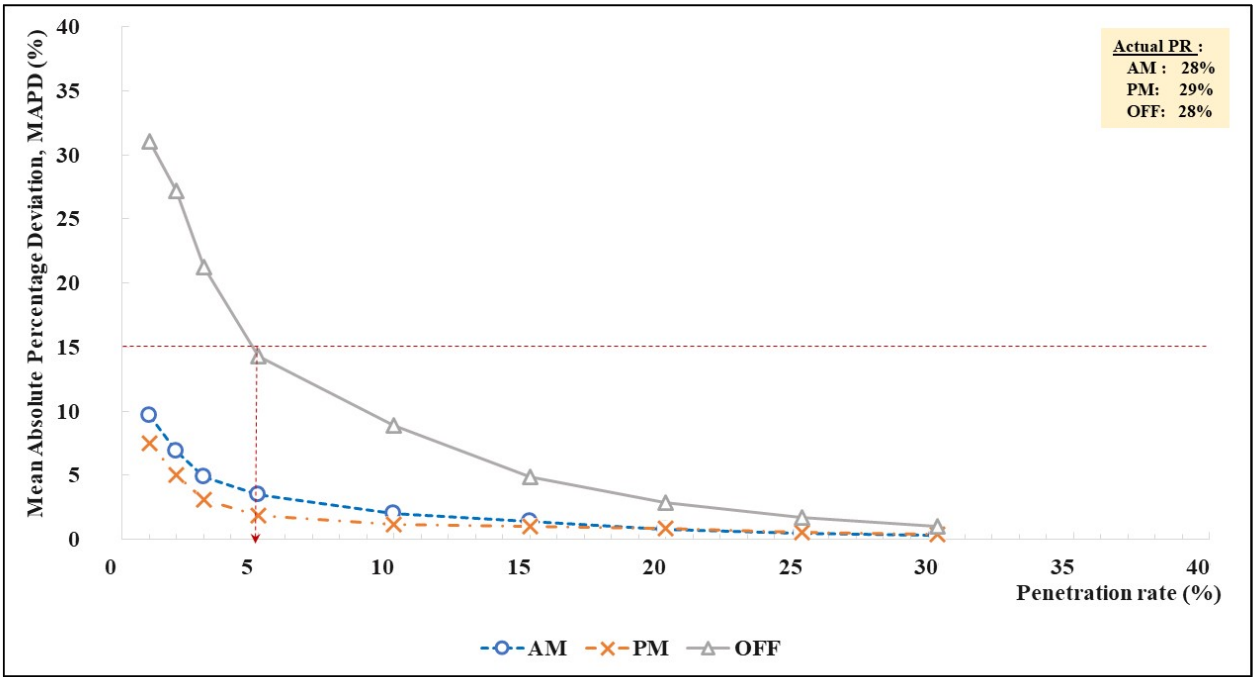

Figure 18. Mean Absolute Percentage Deviation (MAPD) of 50th percentile travel time at different penetration rates for outbound direction on Friday 2 October 2020.

For the outbound direction, the MAPD values for both the AM and PM peaks were lower than the 15 percent threshold for all penetration levels, while that of the off-peak period exceeded the threshold when the penetration rate was lower than 5 percent (equivalent to 49 MAC addresses/h), similar to that of the inbound off-peak traffic.

Comparing travel-time data among the AM and PM peak and the off peak, the off peak tended to have a higher MAPD than the AM and PM peaks at the same penetration rate. This could have been because vehicles were traveling at various speeds during light traffic periods, while during peak hours they traveled at a more uniform speed.

5.6. Impact of Penetration Rate on Estimated 85th Percentile Travel Time

The sensitivity of the 85th travel-time percentile to the penetration rates is shown in Figure 19 (inbound) and Figure 20 (outbound). For the inbound direction, the off peak required a higher penetration rate (5%) to maintain the MAPD at 15 percent or less, followed by the PM peak (medium traffic), and lastly the AM peak (heaviest traffic). For the outbound direction, the off peak also required at least 5 percent of the Bluetooth penetration rate to maintain the MAPD under 15 percent, followed by the AM peak (medium traffic) and lastly the PM peak (heaviest traffic).

The minimum penetration rates as well as the minimum numbers of vehicles/h for both the 50th and 85th percentile travel times are summarized in Table 13. Considering the Bluetooth penetration rates, the heavier traffic scenarios (the AM peak for the inbound traffic and the PM peak for the outbound traffic) required lower penetration rates (1 percent or lower) to maintain travel-time estimation accuracy, while the off-peak periods required a higher percentage (at least five percent) for the Bluetooth penetration rate. However, considering the minimum number of Bluetooth-equipped vehicles/h, the gap between the heavy and light traffic conditions was relatively smaller. For example, the minimum number of Bluetooth-equipped vehicles for the inbound traffic for the AM peak (<26 vehicles/h) was slightly lower than for the off-peak period (34 vehicles/h). The outbound traffic also required 34 vehicles/h for the PM peak, compared with 49 vehicles/h for the off-peak period.

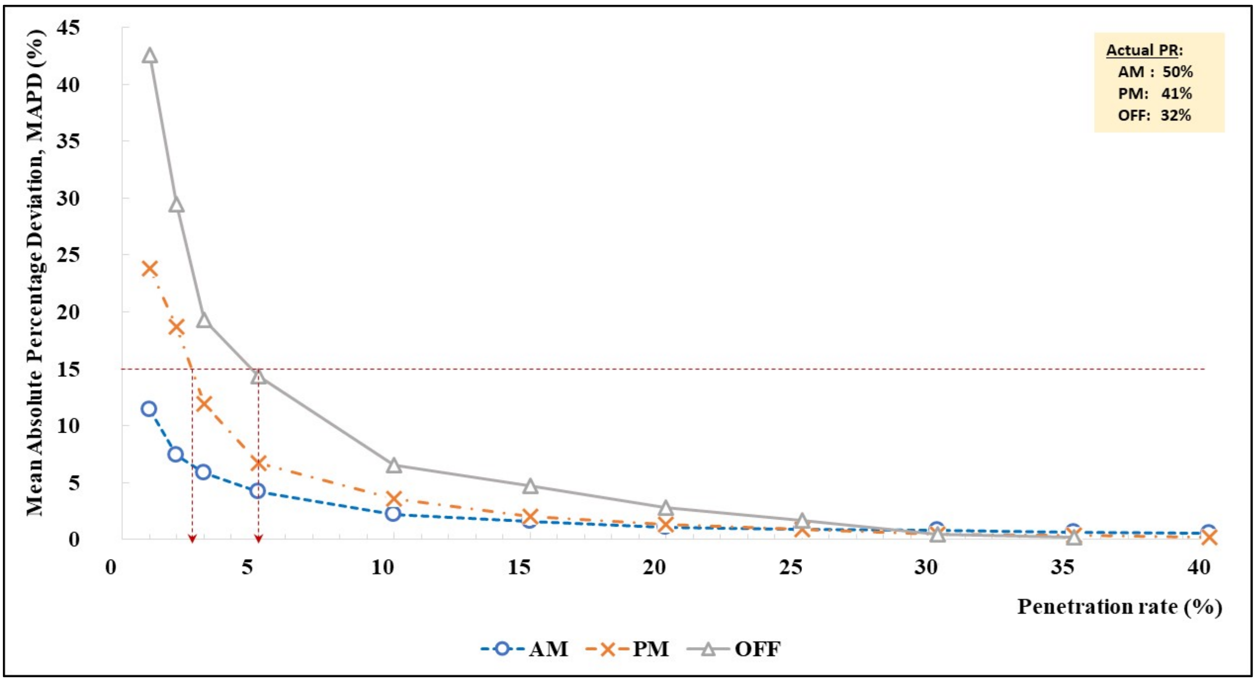

Figure 19. Mean Absolute Percentage Deviation (MAPD) of 85th percentile travel time at different penetration rates for inbound direction on Monday 26 October 2020.

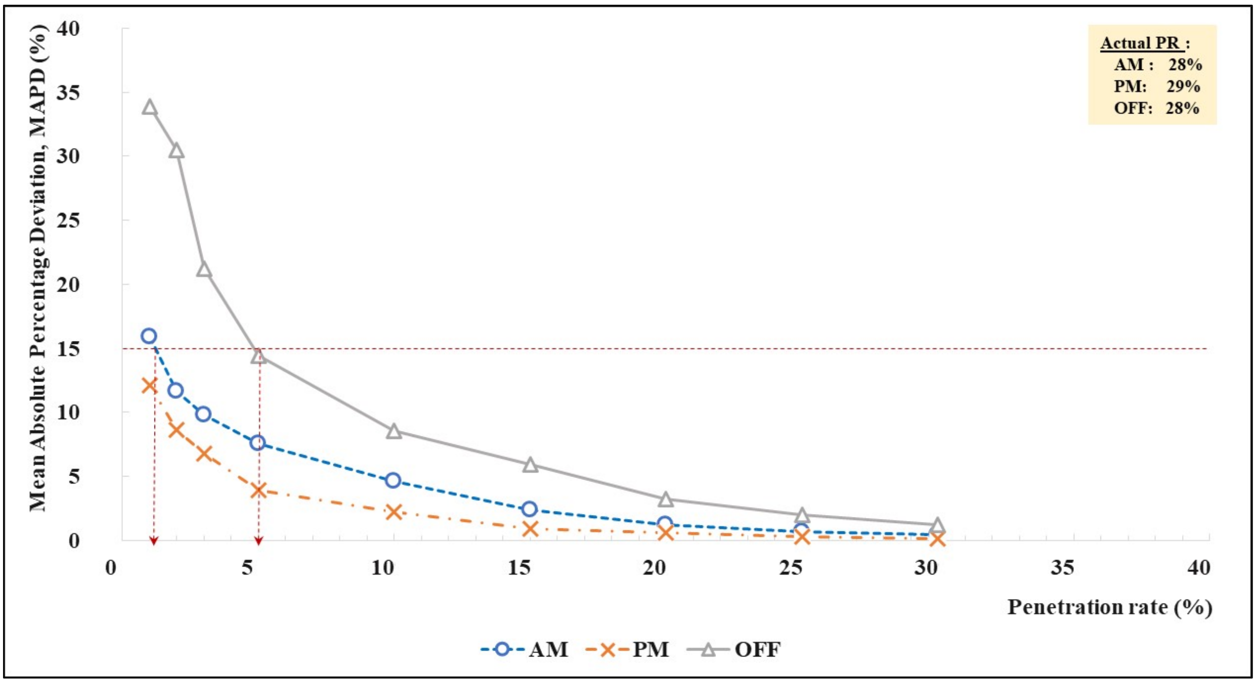

Figure 20. Mean Absolute Percentage Deviation (MAPD) of 85th percentile travel time at different penetration rates for outbound direction on Fridays 2 October 2020.

Table 13. Minimum Bluetooth penetration rates with mean absolute percentage deviation ≤ 15%.

Period	Minimum Bluetooth Penetration Rate (%)			
	Inbound Traffic on Monday 26 October 2020		Outbound Traffic on Friday 2 October 2020	
	50th Percentile Travel Time	85th Percentile Travel Time	50th Percentile Travel Time	85th Percentile Travel Time
AM peak (6:00 a.m.–9:00 a.m.)	<1% (<26)	1% (26)	<1% (<16)	1% (16)
PM peak (4:00 p.m.–7:00 p.m.)	1.5% (37)	3% (56)	<1% (<34)	<1% (<34)
Off peak	5% (34)	5% (34)	5% (49)	5% (49)

Remark: Numbers in parentheses denote minimum number of vehicles/h for each scenario. Light-gray-shaded cells refer to heaviest traffic conditions for the traffic direction (AM peak for inbound traffic and PM peak for outbound traffic).

In addition, the 50th percentile travel time generally required a lower Bluetooth penetration rate and higher numbers of Bluetooth-equipped vehicles than the 85th percentile travel time. Therefore, the 50th percentile statistics are recommended for reporting travel-time information when the Bluetooth penetration rate is low.

The findings from the analyses in this section can be summarized as follows:

- We compared the number of MAC addresses detected by two Bluetooth detectors, where one detector (No. 9) was located on the free-flow segment and the other detector (No. 8) was located 300 m from the toll plaza. We found that the number of MAC addresses detected at Bluetooth No. 8 (slow traffic) was consistently higher than the number detected at Bluetooth No. 9 (free-flow traffic). This implied that Bluetooth performance increased as vehicle speed decreased.
- Next, we analyzed the sensitivity of the estimated travel time to various levels of Bluetooth penetration rates. We found that the mean absolute percentage deviation (MAPD) decreased as the penetration rate increased.
- MAPD during the off-peak period was higher than during the peak period due to the wider range of speed choice during the off-peak period. Therefore, the travel time during the off-peak period required a larger sample size (34–49 vehicles/h) than during the peak period (26–34 vehicles/h) to maintain the same level of MAPD.
- The 50th percentile travel time was less sensitive to various traffic conditions and penetration rates compared to the 85th percentile travel time. Therefore, it is recommended to use the 50th percentile travel-time estimate when traffic conditions are unstable, or the penetration rate is low.

6. Conclusions

This study addressed how Bluetooth penetration rates affected the estimated travel times on a toll road. First, data cleaning and processing were developed to estimate travel times from Bluetooth detector transactions. To obtain Bluetooth trips that could later be used in travel-time estimation, five data-processing steps were developed: cleaning duplicate data; creating a Bluetooth "link" by matching the same MAC addresses detected at two locations; identifying movement direction; determining Bluetooth trip (a series of Bluetooth links); and removing MAC addresses in the same vehicle. As a result, 2 million trips were identified from the 20 days of collected Bluetooth MAC address data.

To determine the Bluetooth penetration rate (the proportion between the detected Bluetooth-equipped vehicles and the total number of vehicles traversing the study corridor), the number of Bluetooth trips were compared with the traffic volume collected by the microwave radar devices installed near the Bluetooth detector stations. The penetration rate during the day was from 50 to 90 percent, while at night, the penetration rate was from 20 to 50 percent; these values were similar to another study [21]. These results could have been due to the fact that Bluetooth detectors perform better when traffic is slower during the day. Therefore, we also determined the number of MAC addresses detected between two Bluetooth stations, from Bluetooth No. 9 (high-speed midblock) to Bluetooth No. 8

(midblock approaching a toll plaza, hence a slower traffic speed). It was found that the number of MAC addresses detected at Bluetooth No. 8 was consistently higher than the number detected at Bluetooth No. 9. Therefore, we concluded that the Bluetooth detection performance increased as the vehicle speed decreased.

Lastly, sensitivity analysis was conducted to determine the sensitivity of the estimated travel times (both 50th percentile and 85th percentile travel times) to various simulated levels of Bluetooth penetration rates. The travel-time values at the highest penetration were assumed to be the actual travel times. The mean absolute percentage deviation (MAPD) was proposed to measure the deviation of the travel-time values when the Bluetooth penetration rate decreased. From the simulated scenarios, we found that the MAPD decreased as the penetration rate increased. Light traffic conditions, such as during the off-peak period, had higher MAPD values than during the AM and PM peaks, since drivers had a wider range of speed choices during the off-peak period compared to the forced flow at a more uniform speed during peak periods.

The Bluetooth penetration rate should be at least 1 percent during the peak period and 5 percent for the off-peak period. Considering number of Bluetooth-equipped vehicles, the peak periods required at least 26–34 vehicles/h while the off-peak period required 34–49 vehicles/h. Additionally, the 50th percentile travel time was more robust than the 85th percentile travel time, indicating that the former required a smaller Bluetooth sample size to maintain the fifteen percent MAPD threshold.

Author Contributions: The authors confirm their contributions to the paper as follows: K.J.: conceptualization, data curation, formal analysis, writing—original draft preparation, funding acquisition, project administration, and resources; S.B.: conceptualization, methodology, investigation, writing—review and editing, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Thailand Research Fund, grant number MRG5580224.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This study was supported by the Thailand Research Fund (Project No. MRG5580224). The authors also thank Don Muang Tollway Public Company Limited for providing support in data collection.

Conflicts of Interest: The authors declare no conflict of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Chen, C.; Ding, Y.; Xie, X.; Zhang, S.; Wang, Z.; Feng, L. TrajCompressor: An Online Map-matching-based Trajectory Compression Framework Leveraging Vehicle Heading Direction and Change. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 2012–2028. [CrossRef]
2. Chen, C.; Jiao, S.; Zhang, S.; Liu, W.; Feng, L.; Wang, Y. TripImputor: Real-Time imputing taxi trip purpose leveraging multi-sourced urban data. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 3292–3304. [CrossRef]
3. Huang, T.; Ma, Y.; Qin, Z.T.; Zheng, J.; Liu, H.X.; Zhu, H.; Ye, J. Origin-destination Flow Prediction with Vehicle Trajectory Data and Semi-supervised Recurrent Neural Network. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 1450–1459. [CrossRef]
4. Tolouei, R.; Psarras, S.; Prince, R. Origin-Destination Trip Matrix Development: Conventional Methods versus Mobile Phone Data. *Transp. Res. Procedia* **2017**, *26*, 39–52. [CrossRef]
5. Fekih, M.; Bellemans, T.; Smoreda, Z.; Bonnel, P.; Furno, A.; Galland, S. A Data-Driven Approach for Origin–Destination Matrix Construction from Cellular Network Signalling Data: A Case Study of Lyon Region (France). *Transportation* **2021**, *48*, 1671–1702. [CrossRef]
6. Cragg, S. Bluetooth Detection-Cheap But Challenging. In Proceedings of the Scottish Transport Applications and Research Conference (STAR), Glasgow, UK, 6 May 2012.
7. Kaur, J.; Kaur, R.; Kaur, M. Bluetooth Technology. *Int. J. Eng. Comput. Sci.* **2016**, *5*, 16001–16003. [CrossRef]
8. Arista.com Santa Clara, CA. 2020, p. 3. Available online: https://www.arista.com (accessed on 26 February 2022).

9. Doherty, P. A review of Bluetooth security in the automotive environment. In Proceedings of the Irish Signals and Systems Conference 2004, Belfast, Ireland, 30 June–2 July 2004; pp. 514–519. [CrossRef]
10. Becker, J.K.; Li, D.; Starobinski, D. Tracking Anonymized Bluetooth Devices. *Proc. Priv. Enhan. Technol.* **2019**, *2019*, 50–65. [CrossRef]
11. Abedi, N.; Bhaskar, A.; Chung, E. Bluetooth and Wi-Fi MAC address based crowd data collection and monitoring: Benefits, challenges and enhancement. *Australasian Transport Research Forum 2013 Proceedings*. Australasian Transport Research Forum 2013. Available online: https://eprints.qut.edu.au/71808/ (accessed on 26 February 2022).
12. Verma, M.; Singh, S.; Kaur, B. An Overview of Bluetooth Technology and its Communication Applications. *Int. J. Curr. Eng. Technol.* **2015**, *5*, 1588–1592.
13. Tsira, V.; Nandi, G. Bluetooth Technology: Security Issues and Its Prevention. *Int. J. Comput. Appl. Technol.* **2014**, *5*, 1833–1837.
14. Araghi, B.N.; Hammershøj Olesen, J.; Krishnan, R.; Tørholm Christensen, L.; Lahrmann, H. Reliability of Bluetooth Technology for Travel Time Estimation. *J. Intell. Transp. Syst. Technol. Plann. Oper.* **2015**, *19*, 240–255. [CrossRef]
15. Tahmasseby, S. Traffic Data: Bluetooth Sensors vs. Crowdsourcing—A Comparative Study to Calculate Travel Time Reliability in Calgary, Alberta, Canada. *J. Traffic Transp. Eng.* **2015**, *3*, 63–79. [CrossRef]
16. Tsubota, T.; Yoshii, T. An Analysis of the Detection Probability of MAC Address from a Moving Bluetooth Device. *Transp. Res. Procedia* **2017**, *21*, 251–256. [CrossRef]
17. Namaki Araghi, B.; Skoven Pedersen, K.; Tørholm Christensen, L.; Krishnan, R.; Lahrmann, H. Accuracy of Travel Time Estimation Using Bluetooth Technology: Case Study Limfjord Tunnel Aalborg. *Int. J. Intell. Transp. Syst. Res.* **2015**, *13*, 166–191. [CrossRef]
18. Friesen, M.R.; McLeod, R.D. Bluetooth in Intelligent Transportation Systems: A Survey. *Int. J. Intell. Transp. Syst. Res.* **2015**, *13*, 143–153. [CrossRef]
19. Sharifi, E.; Hamedi, M.; Haghani, A.; Sadrsadat, H. Analysis of vehicle detection rate for bluetooth traffic sensors: A case study in maryland and delaware. In Proceedings of the 18th World Congress on on Intelligent Transport Systems, Orlando, FL, USA, 10–16 October 2011; Volume 2, pp. 1364–1375.
20. Margreiter, M.; Busch, F.; Carstensen, C. The Evolution of Bluetooth Detection Rates. *Transp. Res. Board* **2019**. Available online: https://mediatum.ub.tum.de/1476803 (accessed on 26 February 2022).
21. Biora, F.; d'Aprile, F.; Marino, R. A Large Scale Application for Bluetooth-Based Travel Time Measurement in The Netherlands. 2012. Available online: https://trid.trb.org/view/1265044 (accessed on 26 February 2022).
22. Cvetek, D.; Bojic, V.; Jelušić, N.; Mustra, M. Initial Bluetooth Probe Vehicle Penetration Rate Analysis: A Case Study in the City of Zagreb. *Int. Sci. Conf. Sci. Traffic Dev. Next Gener. Transp. Ind. Innov.* **2019**. Available online: http://161.53.22.65/datoteka/1000404.Cvetek-Boji-Jelui-Mutra_pross_version.pdf (accessed on 26 February 2022).
23. Ban, X.J.; Li, Y.; Skabardonis, A.; Margulici, J.D. Performance evaluation of travel-time estimation methods for real-time traffic applications. *J. Intell. Transp. Syst. Technol. Plann. Op.* **2010**, *14*, 54–67. [CrossRef]