DOMINIK CVETEK, M.Sc[1]
E-mail: dcvetek@fpz.hr
VEDRAN BOJIĆ, B.Sc[1]
E-mail: 0135226803@fpz.hr
NIKO JELUŠIĆ, Ph.D[1]
E-mail: njelusic@fpz.hr
MARIO MUŠTRA, Ph.D[1]
E-mail: mmustra@fpz.hr
[1] University of Zagreb
Faculty of Transport and Traffic Sciences
Vukelićeva 4, 10000 Zagreb, Croatia

INITIAL BLUETOOTH PROBE VEHICLE PENETRATION RATE ANALYSIS: A CASE STUDY IN THE CITY OF ZAGREB

ABSTRACT

The most of Intelligent Transport System applications rely on various traffic flow data collection technologies. Bluetooth is widely used standard for wireless short-range communication. It is not dedicated for traffic data collection, but Bluetooth devices located in vehicles make them act as detectable probes in traffic flow. That allows collecting of very desirable traffic flow data such as travel times and constructing origin-destination matrices. The Bluetooth probe penetration rate analysis is the first and necessary step for assessment of Bluetooth as traffic flow data collection technology in the area of interest. In this paper we bring field measurement, raw data processing, and filtering with robust filters for detecting outliers in vehicle detection using Bluetooth signal as a marker. In accordance with the reviewed literature, obtained results show that Bluetooth probe penetration rate in the City of Zagreb is sufficient for some ITS applications, and it is stable throughout the week.

KEY WORDS

Intelligent transportation system; Bluetooth technology; Bluetooth vehicle probe; traffic flow data; Bluetooth penetration rate; Vehicle detection;

1. INTRODUCTION

Traffic flow data collection technologies are essential for Intelligent Transportation Systems (ITS) applications. Some of them, which include Bluetooth, can provide very desirable traffic data such as travel time and Origin-Destination (O-D) matrix. Bluetooth (BT) is trademark for short range radio communication standard designed for exchanging data between electronic devices which are frequently and widely used in everyday life. BT based detectors can record unique identifiers of BT devices, Media Access Control (MAC)-addresses. Quality of traffic information derived from collected data by BT detectors depends on number and density of BT probes in traffic flow. BT probe is a vehicle which contains at least one detectable BT device. Therefore, measuring BT probe detection rate in overall flow is a necessary step for assessing the applicability of BT detectors in the area of interest.

Today, data collection technologies for traffic management applications can be sorted in three groups: point sensors, point-to-point sensors and area wide sensors. Technologies which provide vehicle detection at multiple locations as they traverse the network belong to the point-to-point sensors and these are: Automated Vehicle Identification (AVI) Systems, Vehicle identification without driver "cooperation" which include Wi-Fi and BT, and Automatic License plate recognition (ALPR) (1). The application of BT for traffic data collection has some obvious advantages: BT detectors have low cost components, number of BT-enabled devices is increasing, poor weather does not influence detector's performance, and this identification principle is far more acceptable in privacy issue then ALPR.

There are number of papers about BT technology application in traffic which show that BT as a traffic data collection technology has been subject of research in the past decade and still remains topicality. The most frequent papers about BT application in traffic deal with travel time estimation and analyzing its accuracy and reliability, for example (2–5). Some papers extend this topic with application in traffic management or with providing data other than travel time. Authors in (6) deals with BT-based travel time for automatic incident detection. (7) presents integration of BT MAC Scanners data and inductive loop data to obtain travel time and density estimation on urban network. (8) used network of BT Low Energy devices and the received signal strength indicator to detect and classify vehicles by their type: personal car, semi-truck and truck.

Because of BT communication establishment procedure and other factors like signal strength, distance, used antennas and speed of the vehicle in the detection area, not every BT device will be detected by a BT detector. In (4), an experiment was described in which BT devices in vehicles are detected only 80% of the time while passing the BT detector. Quality of information derived from BT data depends on penetration of BT probes in overall flow. In (9) it is concluded that penetration rate should be at least 3% of total vehicle volume and minimum of 20 days of data collection for estimation travel time gives reliable results. Some studies concluded that the minimum percentage of BT penetration rate should be at least 5% (10). Authors in (11) proposed a detection probability model for moving Bluetooth devices. Empirical measurement of BT probe penetration rate is reported in a few studies (12–14). Authors in (12) reported the average rate of 4.5%, in (13) of 29%, and in (14) between a minimum of 4.3% and a maximum of 6.9%. However, the literature is not rich with papers giving precise data about the BT probe penetration rate.

The main objective of this work is to determine the BT penetration rate in overall flow. The rest of this paper is structured as follows. In Section 2 BT standard, test site and used measuring equipment are described. Also, presents proposed filters for outliers. Section 3 combines the case study results and discussion. The final section concludes the case study.

2. METHODOLOGY

Bluetooth standard

BT stands for technology with short-range communication, low-cost components, and low energy consumption. It is convenient for machine-to-machine applications, mobile, voice or stereo handsets, wireless controllers or human interface devices, etc. Bluetooth standard has been constantly evolving since 1994. The last version of Bluetooth is v5.1, presented on 21st January 2019 (15). However, previous versions are still widely used. The commonly used BT versions: Bluetooth v.1.x support data rates up to 1 Mbps, Bluetooth v2.1 + EDR introduced enhanced data rate (EDR) up to 3 Mbps, Bluetooth v3.0+HS 24 Mbps speed rate and Bluetooth v4.0 and Bluetooth Low Energy (15) (16). BT connection between two devices is a multi-step process and those steps are the inquiry, paging or connecting, and connection. It uses the same 2.4-GHz ISM frequency band as RF protocols ZigBee and Wi-Fi (17). The vehicle detection system (VDS) is possible because the BT device has a unique 48-bit address. Usually it is shown in form of a 12-digit hexadecimal value. The half (24 bits) of the address refer to the organization unique identifier (OUI), which identifies the manufacturer. The lower 24-bits are a unique part of the address. Also, important part of BT detection system is its detection range. BT devices are separated into three groups by their power class and range. Class 1 with 100 mW output power and 100 meters range, Class 2 with 2.5 mW output power and 10 meters range, and Class 3 with 1 mW output power and 10 centimeters range (16).

Bluetooth detector

The BT data collection system consists of a BT detector or scanner, a host computer (for gathering and analysing data) and communication, an antenna, power supply, casing, and equipment for mounting. BT is mounted mainly on lighting poles next to roads and it detects passing BT devices, collects data and stores them for future use. Detectors constantly scan the surrounding area in search of an active BT device. Successfully detection means that a device is identified with a unique MAC address. Besides the MAC address, BT detectors receive further information such as a time stamp, name of the device, class, and the strength of the received signal.

The high resolution and time-stamped data can be provided by BT probes to derive traffic metrics such as travel time, O-D matrix, and speed (18). For calculating these measures, two or more BT detectors are required, but there are traffic data that can be provided with only one BT detector.

Number of Hits (N_{Hits}) is the number of times (k) a device or vehicle identified by a unique MAC address is detected in a single detection zone. It is obvious that vehicles or devices are detected more often if they stay longer in the detection zone, based on the fundamentals of the BT detection process (19, 20).

$$N_{Hits} = \sum_{i=1}^{k} MAC\_address_i \ [hits] \tag{1}$$

BT Dwell Time (dwT) is a single detector measurement that is computed from the timestamps of a device entering the detection zone of a detector (first detection) and exiting it (last detection) (19).

$$dwT = Time\ of\ Last\ Hit - Time\ of\ First\ Hit\ [s] \tag{2}$$

Number of Hits and Dwell Time is a surrogate measure for loop detector occupancy. For example, the greater the dwell time, the longer a vehicle "occupies" the detection zone. As it is possible for a device or vehicle to be detected multiple times at a single detection zone, there are several ways to define BT travel time. The difference between the time of detection first – first, first – last, last – last, last – first, average – last, and average – first.

Average BT travel time (TT_{BT}) between nodes A and B during the period P, where n is the number of observations during the period P (21).

$$TT_{BT}(P) = \frac{\sum_{i}^{n_P}\left(t_{i,B} - t_{i,A}\right)}{n_P}\ [s] \tag{3}$$

Data collection and filtering

The result of BT observations are mostly heterogeneous datasets collected from various personal devices used by commuters. (22) used three cluster analysis methods K-means (KM), fuzzy c-means (FCM), and partitioning around medoids (PAM) to automatically classify different travel modes from MAC addresses collected from moving entities. In this paper the goal is to detect only vehicles with BT devices. So, we choose the test site where other modes of transportation are rare. The main idea to classify BT device as "vehicles" relies on recorded BT Dwell Time and Number of Hits. If the BT Dwell Time is reasonable in duration and has a reasonable Number of hits than that device is considered to be a "vehicle". All others are considered as outliers. Detection of outliers is an essential part of data analysis. Outliers may cause a negative effect on data analysis, such as Analysis of variance (ANOVA) and regression, based on distribution of assumptions, or may provide useful information about data when an unusual response to a given study is discovered. Some methods are sensitive to extreme values, and others are resistant to extreme values, like Tukey's and Median rule method (23), used as methods for filtering the outliers. In the section 3 results of these two methods are compared regard to detection rates of BT vehicles.

Bluetooth probe dataset

An example of recorded BT probe dataset is shown in Table 1. Each row in dataset contains columns presenting timestamp in Unix Timestamp format, MAC address, the first part of MAC address, device type in hexadecimal form, and Received Signal Strength Indicator (RSSI). RSSI is measure of signal strength or received signal power in dBm.

Table 1 – Row example in dataset

TIMESTAMP	MAC_ADDRESS	FIRST_MAC	DEVICE_TYPE	RSSI (dBm)
1552565883	48:a9:d2:ef:9c:98	48:a9:d2	340408	-84

BT detector scans the space and if the device is retained in the detection zone for more than 10 seconds, the same MAC address will be recorded. The number of hits is a measure that points to several rows with equal MAC addresses. This leads to a logical conclusion that if devices pass faster through the detection zone, they will have less amount of the number of hits and vice versa (18).

Filtering methods

Tukey's method is applicable to skewed or non-mound shaped data (24). This paper proposes Tukey's method or Interquartile range (IQR) method for filtering BT probe dataset.

In cases when a dataset has outliers or extreme values for describing a dataset by a typical value, the median is used in opposition to the mean. Variability should be summarized by a statistic measure called the interquartile range. The interquartile range is the difference between the first and the third quartile.

The first quartile denoted as $Q1$, represents the value in the dataset that holds 25% of the values below it. The third quartile denoted as $Q3$, represents the value in the dataset that holds 25% of the values above it.

The interquartile range is defined as follows:

$$IQR = Q3 - Q1 \tag{4}$$

Tukey fences filtered data outside the following range:

$$[Q1 - k \cdot IQR, \qquad Q3 + k \cdot IQR] \tag{5}$$

for some non-negative constant k.

Tukey proposed this test, where $k = 1.5$ indicates a less skewed distribution, which is called an inner fence, and where $k = 3$ which indicates data with outliers far out, called an outer fence (24).

The median is a robust estimator of population. It is an 50% breakdown point. The values that are located exactly in the centre of the population when the data is arranged in order (23).

$$Q2 = \begin{cases} \left\{\frac{(n+1)}{2}\right\} th\ value, & if\ n\ is\ odd\ number \\ \frac{\left[\left\{\frac{n}{2}\right\} th\ value + \left\{\frac{n}{2}+1\right\} th\ value\right]}{2}, & if\ n\ is\ an\ even\ number \end{cases} \tag{6}$$

The Median rule *filtered data outside the range*:

$$[-Q2 - 2.3 * IQR, \qquad Q2 + 2.3 * IQR\ ] \tag{7}$$

Measuring equipment and test site

Measuring equipment used in this case study includes BT traffic detector and radar traffic detector. Radar detector is used as ground truth technology for counting the total number of vehicles and measuring their speed. For the power supply we used a solar-powered system independent of outside power source. Photovoltaics system contains a solar panel, battery, and charge controller.

Frequency Modulated Continuous Wave Radar (FMCW) used for overall flow measurement was Huston SpeedLane, working in 24 GHz K-band. This radar is specifically designed for portable or permanent traffic data measurement and collection (25).

The DeepBlue Sensor v2t R-Model by trafficnow® is used as a BT detector. It detects all BT versions in more than 500-meter range with -104 dB receive sensitivity. It is suitable for usage in large urban metropolitan area or at intersections (26). In this paper, the BT detector is used as a point detector on a road segment.

In this experiment we used a flat patch range extender antenna L-com model RE09P with the gain of 8 dBi. It is suitable for both indoor and outdoor applications in the 2.4 GHz ISM band, including IEEE 802.11b and 802.11g, BT and for public wireless hotspot applications (27). (*Figure 1*) shows antenna gain (radiation) pattern.

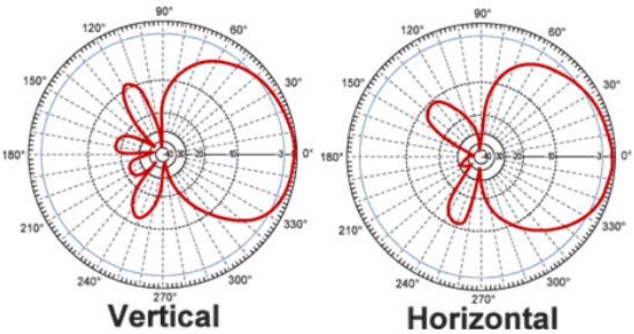

Figure 1 – Antenna gain patterns
Source: (27)

Measurement of BT probe data was done in Zagreb near the underpass - Miramarska street. All equipment was mounted on the concrete pole. Figure 2 shows the exact location of radar and BT detector in Miramarska street.

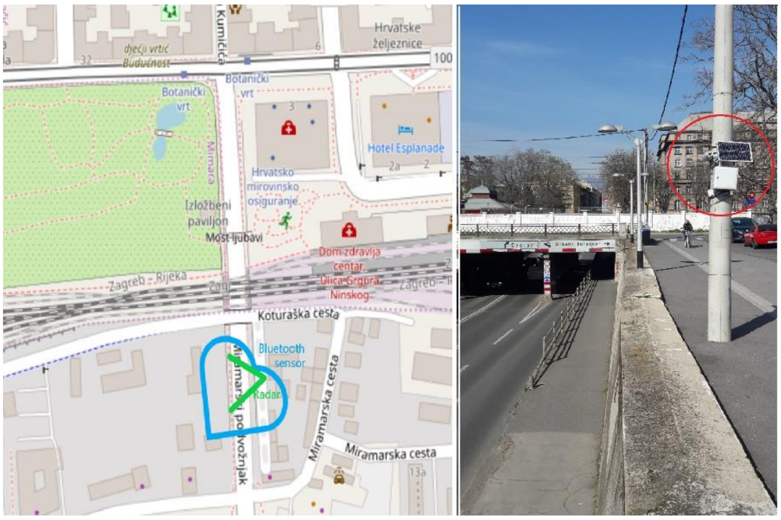

Figure 2 – The location of radar and BT detector

3. RESULTS AND DISCUSSION

BT detector collected 84,767 records between 14th and 15th and between 22nd and 25th March 2019. From 14th March in 15:00 h to 15th March in 15:00 h 4,167 unique MAC addresses, referring to unique devices were recorded and 4,763 from 23rd March in 00:00 h to 26th March in 00:00 h. The observed dataset is used to compare the BT vehicle penetration rate on working and non-working days. Figure 3 shows behaviour of the traffic from 14th March in 15:00 h to 15th March in 15:00 h. *Y*-axis shows that only few devices were detected with large amounts of hits and all three days show a very similar distribution of the number of hits.

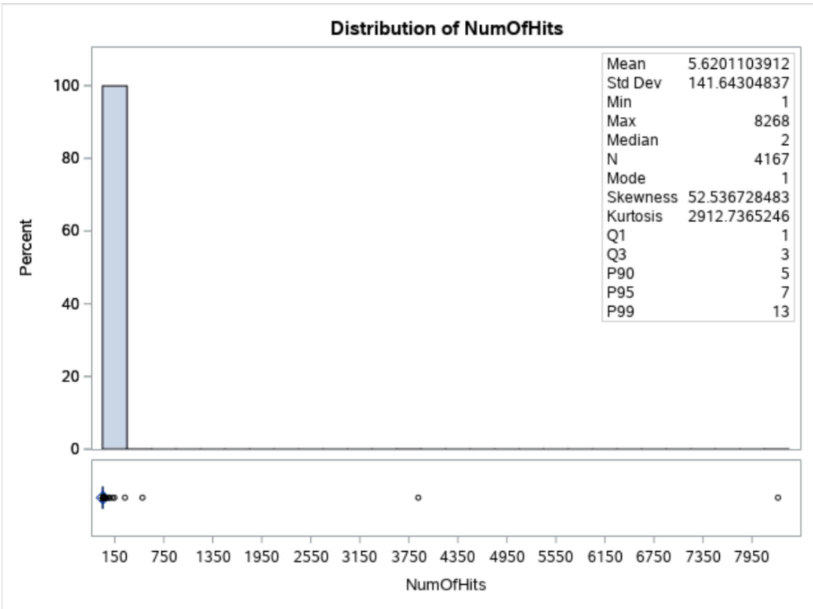

Figure 3 – Frequency of number of hits in observed population

Figure 4 shows distribution when obvious outliers presented by a large value of number of hits are discarded. From frequency distribution in Figure 4, it is evident that most devices have small values of the number of hits.

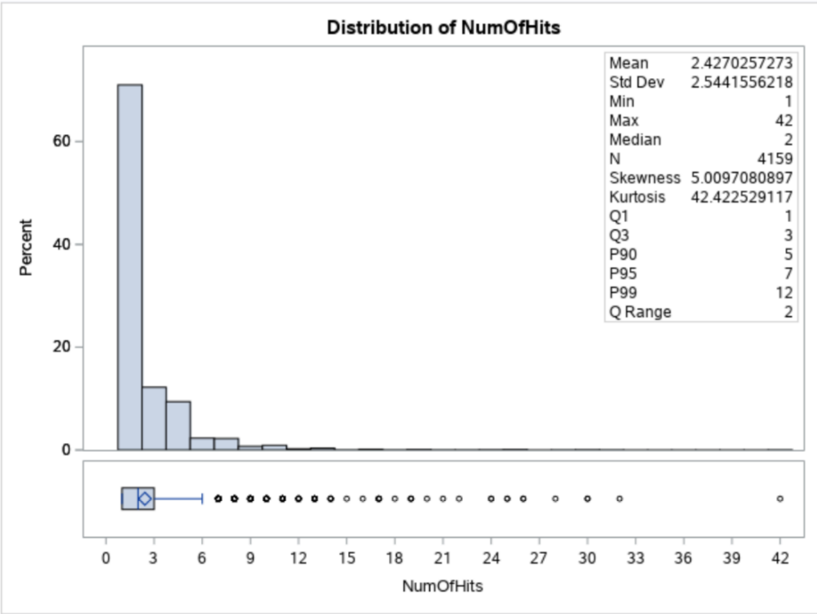

Figure 4 – Distribution of Number of Hits without "tail" values

Collected data show that there is a large dispersion and that points to a premise we are dealing with a heterogeneous set of data. In order to separate vehicles from all recorded BT probe data, filtering was done using the Tukey's method and the Median Rule method.

Table 2 shows data from the Radar detector from Thursday to Friday (24-hour period), Saturday and Sunday.

Table 2 – Radar vehicle data collection

Days	from Thursday to Friday	Saturday	Sunday
Number of vehicles (*N*)	32,379	23,874	17,207

For the weekend period, there is an expected fall in the number of vehicles, but the percentage of vehicles equipped with BT devices remains roughly the same. Figures 5-7 show the comparison between the actual vehicle count measured by the radar detector and the BT share in the actual vehicle count. The BT share for all three days is firstly filtered using the Tukey's method. In Figures 5-7) Blue line (lower line) shows filtered data by the Tukey's method and k=1.5.

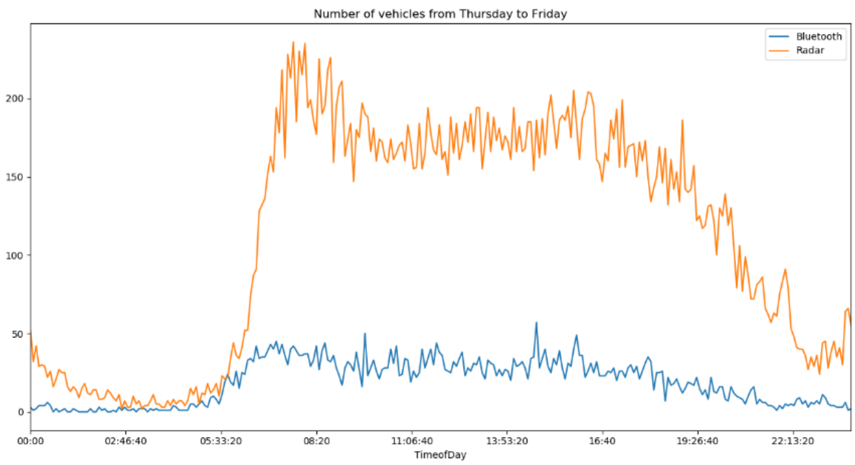

Figure 5 – Radar and BT counts in 5-minute intervals for Thursday and Friday

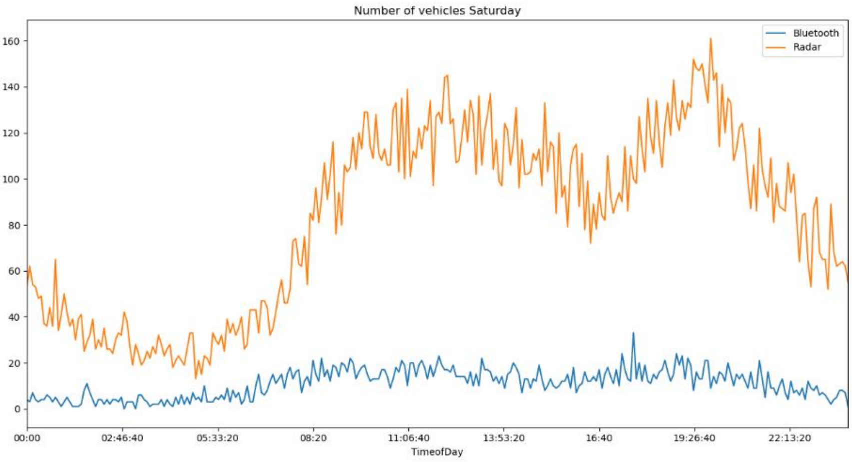

Figure 6 – Radar and BT counts in 5-minute intervals for Saturday

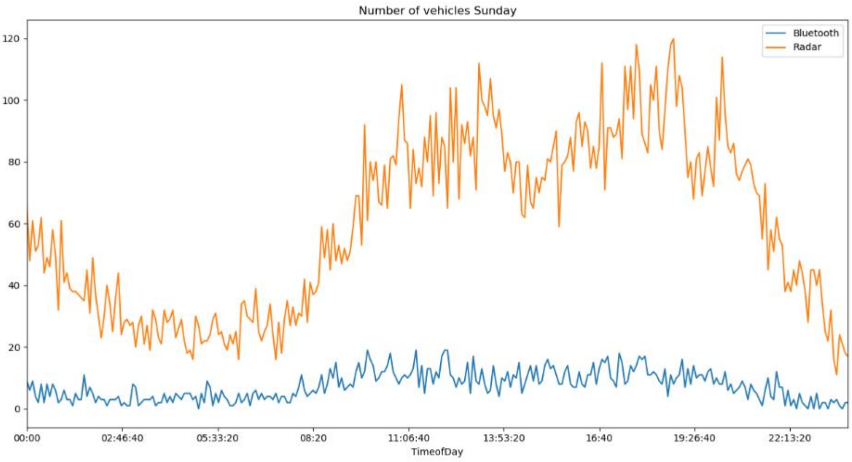

Figure 7 – Radar and BT counts in 5-minute intervals for Sunday

The Tukey's and the Median rule method were applied to data, and the percentage of the vehicles equipped with a BT device is presented in Table 3. The Median rule and the Tukey's fences have the same results for k=1.5. Tukey's method for k=3 against k=1.5 and Median rule have the same difference in percentage, 0.77% (which is for absolute value 183 vehicles for Saturday and 132 vehicles for Sunday). For Thursday to Friday this difference is 0.31% and in absolute value the difference is 97vehicles.

Table 3 – Percentage of vehicles in actual vehicle count

Filtering method	Thursday and Friday	Saturday	Sunday
Tukey's method k=1.5	12.19 %	10.43%	10.00%
Median rule	12.19 %	10.43%	10.00%
Tukey's method k =3	12.5%	11.2%	10.77%

By applying these two methods the percentage of the vehicles equipped with BT devices is in the range between 10% and 12.19%. It can be concluded that BT penetration rate over the time is stable.

The methods used in this paper filters static BT devices, i.e. those with a far greater value of the number of hits, than the rest of the observed population. However, this is not efficient if the measured percentage of BT devices has a large share of cyclists or pedestrians. The number of hits for each MAC address during the observed period is calculated cumulative. This means that the vehicles passing more than once through the road segment have a higher value of the number of hits. This can cause, for example, that the vehicle which passes several times through the detection zone has the same number of hits as the pedestrian that passes through the detection zone once.

However, the Tukey's method for k=1.5 and the Median rule have discarded all consecutive detections of the same devices with the number of hits above 6. This means that devices that have been retained in the detection zone for more than one minute have not been considered in the detection of vehicles equipped with BT devices. it is possible that vehicles that were in congestion were thrown out by this rule. In order to obtain a more accurate estimation of the percentage of vehicles equipped with BT devices, the number of hits should be calculated within intervals and separated by different vehicle routes. It is necessary to find patterns of pedestrian and bicycle behavior to extract these devices from the total number of detected BT devices. To achieve this goal, two or more BT detectors can be used, and cluster analysis methods can be applied to more precisely determine the mode of transport detected from BT detectors.

4. CONCLUSION

Traffic data collection based on BT technology has some notable features: possibility of measuring very desirable traffic parameters such as travel time and O-D matrices, increasing rate of Bluetooth devices in vehicles, more affordable assembly components of Bluetooth traffic detectors and easy mounting of Bluetooth traffic detectors. This makes Bluetooth technology an attractive complementary source of traffic data. Field experiment and data analysis show that the penetration rate of BT probe vehicles is in the range between 10% and 12.19%, and it is stable over the time. This statement is justified because robust filtering of outliers was used. For further analysis, pedestrian and bicycle patterns should be taken into consideration as well as cases of usage of multiple BT devices per vehicle. Results of the case study in the city of Zagreb (Croatia), shows that Bluetooth technology can be considered as an acceptable technology for automatic data collection.

REFERENCES

[1] Antoniou C, Balakrishna R, Koutsopoulos HN. A Synthesis of emerging data collection technologies and their impact on traffic management applications. Vol. 3, European Transport Research Review. 2011. p. 139–48.

[2] Barcelö J, Montero L, Marqués L, Carmona C. Travel Time Forecasting and Dynamic Origin-Destination Estimation for Freeways Based on Bluetooth Traffic Monitoring. Transp Res Rec J Transp Res Board [Internet]. 2010 Jan 1 [cited 2019 Apr 3];2175(1):19–27. Available from: http://journals.sagepub.com/doi/10.3141/2175-03

[3] Bhaskar A, Chung E. Fundamental understanding on the use of Bluetooth scanner as a complementary transport data. Transp Res Part C Emerg Technol [Internet]. 2013 Dec [cited 2019 Apr 3];37:42–72. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0968090X13002027

[4] Araghi BN, Hammershøj Olesen J, Krishnan R, Tørholm Christensen L, Lahrmann H. Reliability of Bluetooth Technology for Travel Time Estimation. J Intell Transp Syst [Internet]. 2015 Jul 3 [cited 2019 Apr 3];19(3):240–55. Available from: https://www.tandfonline.com/doi/full/10.1080/15472450.2013.856727

[5] Vinagre Diaz JJ, Rodriguez Gonzalez AB, Wilby MR. Bluetooth Traffic Monitoring Systems for Travel Time Estimation on Freeways. IEEE Trans Intell Transp Syst [Internet]. 2016 Jan [cited 2019 Apr 3];17(1):123–32. Available from: http://ieeexplore.ieee.org/document/7194827/

[6] Karatsoli M, Margreiter M, Spangler M. Bluetooth-based travel times for automatic incident detection – A systematic description of the characteristics for traffic management purposes.

Transp Res Procedia [Internet]. 2017 Jan 1 [cited 2019 Apr 3];24:204–11. Available from: https://www.sciencedirect.com/science/article/pii/S2352146517303903

[7] Bhaskar A, Tsubota T, Kieu LM, Chung E. Urban traffic state estimation: Fusing point and zone based data. Transp Res Part C Emerg Technol [Internet]. 2014 Nov 1 [cited 2019 Apr 3];48:120–42. Available from: https://www.sciencedirect.com/science/article/pii/S0968090X14002319

[8] Bernas M, Płaczek B, Korski W. Wireless Network with Bluetooth Low Energy Beacons for Vehicle Detection and Classification. In Springer, Cham; 2018 [cited 2019 Apr 3]. p. 429–44. Available from: http://link.springer.com/10.1007/978-3-319-92459-5_34

[9] Shahram Tahmasseby. Traffic Data: Bluetooth Sensors vs. Crowdsourcing—A Comparative Study to Calculate Travel Time Reliability in Calgary, Alberta, Canada. J Traffic Transp Eng [Internet]. 2015 Feb 28 [cited 2019 Apr 3];3(2). Available from: http://www.davidpublisher.org/index.php/Home/Article/index?id=5771.html

[10] Maxwell S. FDOT Standard Specification for Road and Bridge Construction. Florida Department of Transportation; 2019.

[11] Tsubota T, Yoshii T. An Analysis of the Detection Probability of MAC Address from a Moving Bluetooth Device. Transp Res Procedia [Internet]. 2017 Jan 1 [cited 2019 Apr 3];21:251–6. Available from: https://www.sciencedirect.com/science/article/pii/S2352146517302314

[12] Friesen MR, McLeod RD. Bluetooth in Intelligent Transportation Systems: A Survey. Int J Intell Transp Syst Res [Internet]. 2015 Sep 29 [cited 2019 Apr 3];13(3):143–53. Available from: http://link.springer.com/10.1007/s13177-014-0092-1

[13] Araghi BN, Pedersen KS, Christensen LT, Krishnan R, Lahrmann H. Accuracy of Travel Time Estimation using Bluetooth Technology: Case Study Limfjord Tunnel Aalborg. ITS World Congr [Internet]. 2012 [cited 2019 Apr 3]; Available from: http://vbn.aau.dk/en/publications/accuracy-of-travel-time-estimation-using-bluetooth-technology(95cacf99-5ccd-4610-984e-ade8daf87f51).html

[14] Sharifi E, Hamedi M, Haghani A, Sadrsadat H. Analysis of Vehicle Detection Rate for Bluetooth Traffic Sensors: A Case Study in Maryland and Delaware. 18th World Congr Intell Transp Syst [Internet]. 2011;(June):1–12. Available from: http://trid.trb.org/view.aspx?id=1240324

[15] Bluetooth Special Interest Group (SIG). Bluetooth Core Specification Version 5.0. Bluetooth Core Specif [Internet]. 2016;2684. Available from: https://www.bluetooth.org/en-us/specification/adopted-specifications

[16] Specifications | Bluetooth Technology Website [Internet]. [cited 2019 Apr 3]. Available from: https://www.bluetooth.com/specifications

[17] Stallings W, Beard C. Wireless communication networks and systems. Pearson Higher Education, Inc.; 2016. 587 p.

[18] Ayodele EG. Using Bluetooth to Estimate Traffic Metrics for Traffic Management Applications [Internet]. School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, UK; 2017. Available from: http://www.albayan.ae

[19] Hart-Bishop J. Advenaced Signal Control Using Bluetooth Detectors [Internet]. University of Waterloo; 2018. Available from: http://hdl.handle.net/10012/13037

[20] Ingegneria F, Di M-G. Recognition of Bluetooth Signals Based on Feature Detection. University of Rome; 2010.

[21] Wang Y, Malinovskiy Y, Wu Y-J, Lee UK. Error Modeling and Analysis for Travel Time Data Obtained from Bluetooth MAC Address Matching. Washington; 2011.

[22] Bathaee N, Mohseni A, Park S, Porter JD, Kim DS. A cluster analysis approach for differentiating transportation modes using Bluetooth sensor data. J Intell Transp Syst [Internet]. 2018 Jul 4 [cited 2019 Mar 25];22(4):353–64. Available from: https://www.tandfonline.com/doi/full/10.1080/15472450.2018.1457444

[23] Seo S. A Review and Comparison of Methods for Detecting Outliers in Univariate Data Sets [Internet]. University of Pittsburgh; 2006. Available from: http://d-scholarship.pitt.edu/id/eprint/7948

[24] Tukey JW. Exploratory data analysis. Reading, Mass.: Addison-Wesley Pub. Co.; 1977.

[25] Huston RADAR [Internet]. [cited 2019 Apr 2]. Available from: www.houston-radar.com

[26] DeepBlue R-Model [Internet]. [cited 2019 Apr 1]. Available from: http://deepbluesensor.com/portfolio/deepblue-r-model/

[27] www.L-com.com [Internet]. [cited 2019 Jan 4]. Available from: https://www.l-com.com/multimedia/datasheets/DS_RE09P-XX.PDF