Origin-destination trip pattern representative by using bluetooth data collection on the arterial road network, Bangkok, Thailand

Rattanaporn Kasemsri\textsuperscript{a,*}, Tetsuhiro Ishizaka\textsuperscript{b}, Atsushi Fukuda\textsuperscript{b}, Piyapong Suwanno\textsuperscript{c}

\textsuperscript{a}School of Civil Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima, 30000, Thailand.
\textsuperscript{b}College of Science and Technology, Nihon University, Chiba, 274-8501, Japan.
\textsuperscript{c}College of Industrial Technology and Management, Rajamangala University of Technology Srivijaya, Nakhon Si Thammarat 80210, Thailand.

(Communicated by Madjid Eshaghi Gordji)

Abstract

In order to properly control traffic, it is necessary to grasp detailed traffic states within a certain area. However, the existing traffic detection methods have high production cost, and detailed traffic conditions still cannot be determined based on those methods. Thus, this study tries to extract detailed traffic states using data from Bluetooth detectors collected over a long period. At first, we observe the variability in origin-destination (OD) trip patterns based on weekdays and weekends, and the time of day (morning peak hour, evening peak hour, and off-peak) in the downtown of Bangkok. Secondly, the traffic speed pattern is extracted based on the detected Bluetooth data. Finally, we estimated change of route choice by tracking the same vehicle using the MAC address during the detection period. The experimental results confirm that the detailed traffic states such as OD trip pattern, speed pattern and route choice behavior could be observed by Bluetooth data which could not be observed by other traffic data measurement.

Keywords: Bluetooth data, Traffic state, OD pattern, Route choice behavior, bangkok

1. Introduction

Many cities in Southeast Asia do not have adequate vehicle sensors for grasping traffic conditions because of which the traffic is not appropriately controlled by traffic signals. Under such circumstances, traffic congestion has become more acute as the use of automobiles has increased rapidly. In

\textsuperscript{*}Corresponding Author: Rattanaporn Kasemsri
Email addresses: rattanaporn7777@gmail.com (Rattanaporn Kasemsri\textsuperscript{a,*}), iishizaka@nihon-u.ac.jp (Tetsuhiro Ishizaka\textsuperscript{b}), fukuda.atsushi@nihon-u.ac.jp (Atsushi Fukuda\textsuperscript{b}), piyapong.su@rmutsy.ac.th (Piyapong Suwanno\textsuperscript{c})

Received: March 2021 Accepted: June 2021
order to control traffic signals appropriately in correspondence with traffic conditions, it is necessary to grasp traffic conditions.

For the cities in developed countries, many vehicle sensors were installed already such as ultrasonic sensors, infrared sensors, inductive loop detectors and etc. so that traffic states can be observed by them. However, introducing these sensors into developing cities is expensive and requires significant installation time which imposes additional limitations on the deployment. Also, it is very difficult to properly install a vehicle sensor in a city that is expanding continuously along with changing road situations. The city of Bangkok in Thailand has one of the highest traffic congestion rates in the world with ever expanding road networks and increasing traffic demand. Further, there are no proper traffic detection systems in Bangkok which can cover whole metropolis. The existing traffic measurement techniques like Taxi-probe, Bus-probe, and Truck-GPS system are not effective to measure the traffic behavior of the majority of vehicles in the city.

Thus, instead of the existing systems, attention has been drawn to a Bluetooth detection method which is able to grasp the traffic situation using information obtained from commonly used devices such as a cellular phone, hands-free headsets, global positioning system (GPS) units and computers. Bluetooth technology is a cost efficient method and is easy to install compared to the other sensors, and allows observing traffic states continuously. An integrated Bluetooth system on vehicles makes the remote detection of these devices possible. And, it makes possible capture and identifies anonymously a significant portion of the traffic stream at a relatively low cost so that this approach becomes increasingly popular. Moreover, it can be used to observe the travel route of vehicles by detecting Media Access Control (MAC) addresses and time stamps at multiple places. Obtaining the unique Mac address provides the opportunity to track behavior and travelling schedule of the same device unlike other sensor systems. Therefore, in this study, we aim to examine applicability of Bluetooth detectors for detection of traffic state in downtown Bangkok.

2. Literature review

The concept of using Bluetooth technology has two components which are MAC address and detected time. The detected time of the same Mac address is sorted in an increasing order so that we can determine the difference between the detected times which provides the travel time of that vehicle. According to this concept we can conclude that the Bluetooth scanner can be used to collect travel time when traveling from one intersection to another intersection in both directions. These data can be scoped by the date, MAC address and detection by detector, respectively. For those MAC addresses which are detected only by a single detector in a day, the trip could not be recorded and the corresponding data is removed. On the other hand, the MAC addresses which are detected by multiple detectors in a single day could reveal the travel paths taken during different trips. Previous studies (Charle et al., 2010; Araghi et al., 2012; Zhang and Haghani, 2015; Vanajakshi et al., 2009) have estimated the route travel time and their accuracy from either real-time information or historical link travel time observations. Moreover, an integrated Bluetooth system on vehicles has made remote detection of these devices an increasingly popular method of capturing and anonymously identifying a significant portion of the traffic stream at a relatively low cost. It is used to observe the travel route of vehicles by detecting MAC addresses and time stamps which are provided by Bluetooth signal at many places (Schneider, et al. 2010; Slone, 2011; Young, 2008; Puckett and Vickich, 2010; KMJ Consulting 2010; Porter et al., 2010; Martchouk and Mannering, 2009; Tarnoff et al., 2009; Barceló et al., 2009).

Reliability of Bluetooth detection is deeply depending on detection range. Thus, selection of the type of antenna is very important to confirm detection range. Porter et al. (2013) have tested five
different types of antennas. They first characterized to understand how their main properties (i.e., polarization and gain) influence the performance BT reader. Those antennas were used to collect travel time samples on a 0.75 mile four-lane road segment in Salem, Oregon. The results indicate that the quantity and the quality (e.g., percentage of duplicates) of the MAC addresses read with each antenna type are different. The omnidirectional antenna, which was indicated has significantly higher fraction read than any other antenna at rate 0.109 while the others are 0.097, 0.090, 0.090 and 0.082 respectively. Other studies have also been performed to evaluate omnidirectional versus directional antennas (Malinovskiy et al., 2011; Malinovskiy et al., 2010; Wang et al., 2010) and the height of the antenna above the road (Brennan et al., 2010). Regarding to previous studies, our study is using this type of antenna, omnidirectional antenna.

Another problem regards on the performance of Bluetooth detection, detection rate. When large numbers of Bluetooth devices are existed in a traffic stream, there is some limitation to detect all of them so that it is very hard to cover all those MAC addresses before they leave the detection range. Thus, using multiple Bluetooth readers in one location has the potential to increase the overall detection rate. If the increase in number of detections is significant, then the benefit of collecting additional data may exceed the additional cost of installing multiple readers (The Bluetooth Special Interest Group, 2010; Woodings et al., 2002).

3. Objective, study area and data

3.1. Objective of study

Traffic data is essential for proper traffic control and those traffic data were collected using traffic detectors. Unfortunately, these detectors are expensive to install so that it is hard to install appropriately. In the case of Bangkok, traffic detectors had been installed in the limited area and traffic state can’t observe properly in whole city area. Even floating car system has been operated recently, data is quite limited. Thus, there is no data to control traffic in real time in Bangkok. Since it is expected that traffic data can be collected the large volume of data with relatively low cost by applying Bluetooth technology, the problem above mentioned can be solved. In addition, by adapting the Bluetooth technology, changes in demand by temporary changes in the network, including closing of roads, unusual weather conditions, or special events can be detected. Thus, this study examined the approach to extract traffic data from Bluetooth data collected for long period.

The objectives of this study are:

1) To extract Bluetooth data on experience assumptions and analyze origin destination trip patterns (OD patterns) of the huge possible data.

2) To recover the missing links for overview speed estimation in study area, Bangkok.

3) To classify the trip frequency by individual unique MAC address and track those trips to see route choices behavior.

3.2. Study area

Bluetooth detectors were installed at 44 main intersections in Bangkok, those locations are shown in figure[1] Their names are represented by device ID. With cooperation of the police, those detectors were set inside of the police box to avoid heat. Previous studies have examined the methods to increase sample size through Bluetooth reader by placing them on medians or on either side of the road (Malinovskiy et al., 2011; Brennan et al., 2010).
3.3. Data

The data utilized in this study is collected by the collaborative research between Chulalongkorn University (CU), Nihon University (NU) and other universities. Since, the Bluetooth detectors were developed by embedding with the 4G communication module and omnidirectional antenna to the raspberry pi computer board. Also, it was developed as own Bluetooth scanning device by using ship processing. Total 44 detectors were installed in October 2017 and they were activated completely in December 2017. The experiment to detect Mac address from installed detectors along the arterial road network in the central city of Bangkok was carried out. On the first trial, the Bluetooth data were detected totally 16 days from December 29, 2017 to January 13, 2018. Then, detection period was expanded and data was detected approximately five months from January 2018 to April 2018. In this paper, this data was analyzed.

4. Extracted bluetooth data

Due to the technology itself which provide the short-range detectable data also the environmental state which is always fluctuate and affect to the detectors. As the results of the low sampling rate obtained by Bluetooth data is the critical issue for reliability of the model. Moreover, the mac address was detected time from origin to destination but some parts in the route detected by Bluetooth is missing lead to the intra-route data possibly imprecise or even unknown. The missing link is may strongly effects for estimation in arterial urban road network unlike freeway that have been study in the past. Therefore, this study is seeking to preliminary preparing Bluetooth data with reliability. The shortest path algorithm is used for searching the missing link in this study. This method is proper to acknowledge about the problem that helps direct search to more promising paths. This information can significantly relate the decision making on both provider and user of traffic planning. Therefore, the missing link needs to be completed by shortest path algorithm beforehand.

4.1. Extracted bluetooth data

The initial data were stored in the main server that can be access with security code. The data can be exported as csv file by date. Figure 2 is the raw data that we have downloaded from the server. Each column has their own meaning following by these explanations. Column V1 is row
number of Bluetooth data on this database name “Tracking”, which has no significant to analyze our data. Column V2 has no meaning. Column V3 is the MAC address from Bluetooth detection. This is the first key value that we will use for analysis. Column V4 means for gain of an antenna system relative to an isotropic radiator at radio frequencies which is not relate to our data analysis. Column V5 is date and time that MAC address was detected by Bluetooth detector, time is in unit UTC +0.00. This is the later key after Mac address that we will use to sort the data and get travel time. Column V6 is date and time data that has send to Server data base. However, we do not use data on this column to calculate travel time data. The last key column V7 stands for ID number of Bluetooth detector. In other word, this column means the location of those Mac address were detected which is necessary to know for analysis. The conceptual framework to extract useable data from raw data will be discussed in section 4 Screening Data Outlier.

According to the detector technology, the data were collected and store via the server. It is necessary to understand the structure of detection data in primary step see in figure 3. Generally, due to the short-range detection and weakness signal technology itself highly possible cause to lack of consecutively data or even precise data. This section shows the tracking way how to extract the
useable data from raw data which is start from the server export. Date will be the first attribute that use to classify the trip as we assume that the trips were detected by day. Mac address comes later for classification those trips in that day. In the same category of unique Mac address will be filtered by number of detector detection. The detector that is detected only once implies that they cannot represent as the trips, the trips must be generated at least from two of detectors. In this primary data filtering, we remove those single detector detections and store multiple detector detections represented as travel paths.

4.2. Validation of bluetooth data

For the validation of extracted Bluetooth data, traffic volume and travel time collected by Bluetooth data are compared with actual collected by the other data source. The traffic volume with same date when Bluetooth data was collected is not available so that the traffic volume data investigated by the Bangkok Metropolitan Agency was employed in this study. This data is intersection traffic volume data for 12 hours. The ratio of actual traffic volume and unique mac address is compared over 15 intersections as shown in figure 4. The average ratio is 4.8 percentage and its validation seems to be a little bit high because it detection ratio may be depending on coverage area of Bluetooth scanner. In some intersections, there are some obstructions such as pier of flyover. But this ratio on each intersection is stable to traffic volume because a condition of obstruction is not changed dynamically.

In figure 5, the travel time is compared with the collected travel time by video camera shooting. It’s shown that average travel time by Bluetooth data is fitted on average travel time estimated by video camera. Based on the two comparisons on traffic volume and travel time, it’s concluded that the Bluetooth data is able to express a perspective of actual traffic flow in Bangkok. And then, it’s inferred that the estimated traffic condition on the next section partially has reproducibility of actual traffic.

![Figure 4: Traffic volume V.S. Unique Mac address](image)

5. Screening data outlier

The initial step is to filter the outlier that possibly strongly influence with the results. Firstly, examined detector should be considered. After receiving travel paths from the sorted raw data, we
extract the OD trip data, origin nodes and destination nodes were represented from detectors. We assumed that those OD trips which have less detection trips are improperly working and should be removed. This step can remove the error caused by ability of detector. Secondly is to see the overall detection rate, in this study divides the number of detection and mac address by weekly. This step can see the trend of detection rate by date, the data will be removed if there is less detection on those weeks. Finally, speed will be the one of criterions to verify the error detections or the unusual situation such as car stopping or pedestrian walking. Those speed less than 4 km/hour are assumed as unusual trips and will be removed from our analysis. Likewise, the speed over than the usual experience 100 km/hour or strongly high is defined as error detection also removed.

5.1. Detector’s operation

As mentioned above, detector’s operation can be checked by the OD table summary. Figure 6 is the OD table that we summarize from the travel paths in primary step. There are three detectors which are 213, 218 and 236 detect very less detection amount. Most of their detection amount is just about between 0-300 detections. Moreover, one detector which is 246, clarity shows less detection only zero and one detection. However, these two detectors 213 and 246 was not appeared the information of location from the beginning so it was not accounted to our original detection map. Additionally, 14 detectors which are 102, 201, 203, 205, 207, 210, 217, 219, 221, 222, 225, 226, 227 and 234 were not found in our primary data. Therefore, those detectors will be removed and updated remain 29 detector locations in the figure 7a.

Since we have removed four detectors in figure 6a, the total number of detectors from the total 44 detectors remains 29 detectors. There are a few detectors remained in the location out of connected area with the other detectors. These detectors are 231, 240 and 242. We assumed to cut these
three detectors because they cannot tell exactly paths they use. Finally, the detectors are remaining updated as 26 detectors as shown in figure 7.

Figure 6: 6OD table receiving from primary data

Figure 7: Updated Bluetooth detector locations

5.2. Detection amount by period

Figure 8 shows the amount of detection that is proportion by single detection and multiple detections as travel paths or travel routes by week. Similarly, figure 9 shows the amount of unique Mac address that is proportion by single detection and multiple detections. These two pictures show the same trend of detection amount which last three weeks were detected as abnormal detection less amount. Therefore, last three weeks will be removed as unconsidered data. These data obviously see that the detection amount has slightly drop on week 14th and 15th, which is correspondingly with national holidays.
5.3. Speed outlier

Speed is our one criterion to examine the outlier. The speed can be indirectly from Bluetooth data by calculation of travel distance divided by travel time as shown figure 10. As the Bluetooth data can be collected the time stamp that we can calculate for travel time. The distance can be measured directly from the field site study area. However, as mentioned before about short range detection cause less detection rate and missing link data. Therefore, the distance of missing data cannot be measured directly from the field site. We overcome this problem using recovering missing link by BFS’s shortest path which is an algorithm to search the shortest path in a graph from a single node to a single node that we have tested for the experiment data previously.
Absolutely seen at above of thousand \textit{km/hour} of speed is impossible to drive. This outlier can be happened because of inaccuracy of detection technology itself which is an error of time stamp detection lead to excessive short travel time such as one second for those travel paths. Therefore, our assumption is to remove those trips and keep the trips between speeds from 4 to 100 \textit{km/hour}. Moreover, we also cut some mac address unique which state for unidentified such as 00 : 00 : 00 : 00 : 00 : 00.

Missing link is an absent link member that needed to complete a series or resolve a problem. For example, figure 10 is the trip from origin A to destination G. In fact, the A-G trip travel by route A-B-D-E-G. However, the Bluetooth can detect that mac address only from A-B and then E-G. The possible route is probably will should A-C-D-E-G instead of real route A-B-D-E-G if we consider only the shortest path problem. Thus, this study sets the objective function is to choose the link pair in those O-D pair which has the summation as close as possible with the detected data, equation (1). $D_{ij}^k$ is the summation of distance sub link pair $d_{ij}^k$ on trip $k$. Where $i \in N$, $j \in N$, and $N$ is the set of nodes. $k$ is the possible trip $i$ to $j$ from BFS’s shortest path.

$$D_{ij}^k = \sum_{k=1}^{k} \sum_{j=1}^{N} \sum_{i=1}^{N} d_{ij}^k \quad (1)$$

where,

$D_{ij}^k$ : travel distance from all possible trip $k$ from $i$ to $j$ by BFS estimation.

$d_{ij}^k$ : travel distance sub link pair on trip $k$.

![Figure 11: An example of the missing link and BFS traverses a graph](image)
This study uses the Breadth First Search (BFS) algorithm for completing the missing links in missing trips. It is the algorithm that traverses a graph in a breadthward motion and uses a queue to remember to get the next vertex to start a search. While using certain graph algorithms, it must be ensured that each vertex of the graph is visited exactly once. The traverse from node to the graph in 11.

The defining characteristic of this search is that, whenever BFS examines a maze cell c, it adds the neighbors of c to a set of cells which it will examine later. BFS maintains a queue of cells which have been visited but not yet examined (an examination of a cell c consists of visiting all its neighbors). Thus, a cell can have three states:

1) Unvisited: The cell has not yet been visited by BFS.
2) Visited but not examined: The cell has been discovered but BFS has not evaluated whether its neighbors should be added to the queue.
3) Examined: The cell has not yet been visited and all its neighbors are/have been in the queue (for example they are already “Visited but not Examined” or “Examined”).

This method is proper to acknowledge about the problem that helps direct search to more promising paths. The ended result is that BFS will visit all the cells in order of their distance from the entrance. The order in which the vertices are visited are important and may depend upon the algorithm or question that you are solving. First, it visits all locations one step away, then it visits all locations that are two steps away, and so on, until an exit is found. Because of this, BFS has the nice property that it will naturally discover the shortest route through the maze. It employs the following rules and represented by figure 12.

1) Rule 1: Visit the adjacent unvisited vertex. Mark it as visited. Display it. Insert it * in a queue.
2) Rule 2: If no adjacent vertex is found, remove the first vertex from the queue.
3) Rule 3: Repeat Rule 1 and Rule 2 until the queue is empty.
6. Data analysis results

6.1. OD Pattern

After screening data from three criterions in the previous section, figure 13 is a plot for number of trips depending on time of day. Indeed, that the amount of detection trip is low from midnight until 5 am which is comparable trend with the real traffic situation. The detected link can be maximum count during this period approximately 200,000 links which is around 50 percent from the peak during the time of day. Moreover, the trips can be detected dramatically increasing from 6 am until 9 am as it represents for peak hour morning in the usual traffic situation. After that, it tends to slightly decrease during day time from 9 am to 4 pm. Stability remain high detection amount at 4 pm to 6 pm as peak hour in the evening then coherently drop again after 6 pm as off-peak hour.

![Figure 13: The number of link count trips by time of day](image)

The OD patterns which are our main objective to receive from Bluetooth data are shown from figure 16 to figure 14. figure 16(a) to figure 16(d) represent OD pattern for every day, every day morning peak hour, every day evening peak hour and every day off-peak hour respectively. Then, every weekday and weekday by time of day were shown from figure 15(a) to figure 15(d). Finally, figure 14 demonstrate for OD pattern of every weekend.

![Figure 14: Weekend OD patterns](image)
Figure 15: Weekday OD patterns

Figure 16: OD patterns of every day
From OD patterns, we found that the traffics were obviously generated on weekday morning peak hour. Their OD patterns of all day and on weekday morning peak hour were accordingly relative. The consecutive OD pairs which are 243-214, 214-245, 245-244, 244-202 and 202-235 were relatively high travel on weekday morning peak, maximum is approximately 12,000 trips. Therefore, this area possibly trends to congested traffic in widespread area. Remaining of some OD pairs such as 233-108, 233-237, and 233-232 are sporadic heavy trips. The effected on those OD pairs from morning peak hour plus trips generated from work place to home lead to accumulated traffic during evening peak hour traffic similarly with morning peak hour approximately 12,000 trips. However, the critical OD is on the weekday excluding morning peak hour and evening peak hour which is clump along 214-245, 245-244, 244-202 and 202-235, especially OD pair 245-244 generate approximately from 30,000 to 40,000 trips. Therefore, the traffic between this OD pair is critically congested and possibly affected to the adjacent area. While the OD patterns on weekend demonstrated light OD trip patterns for all day of all OD pairs. The OD patterns have ability to define where the traffic planner should play more attention for traffic controls.

6.2. Speed

Nevertheless, not only the OD pattern above mentioned can be seen from Bluetooth data collection, some of facet information can be received from these data. One of this information is speed frequency and average speed in our study area, central of Bangkok. Highest frequency speed that mostly presented in figure 17 is 7.5 to 12.5 km/hour while the average speed is approximately 20 km/hour. The speed frequency decrease against when the speed becomes higher as it is indeed that this area is the arterial road and congestion area that cannot drive with high speed.

6.3. Mac address frequency

Specific information is the frequency classified by unique Mac address as shown in Figure 19. We have gathered the Mac address which was detected from minimum one day detection until the
maximum 102 of detection days. The highest bar is the Mac address were detected only one day approximately 200,000 unique Mac address. However, it is significant seen that there are more than half of all unique Mac address can be detected more than two days or multiple days. This information implies that the same unique Mac address travel in this study area as usually depending on number of detection days. The lowest number of unique Mac address in the highest number of detection days presents for the highest frequency travelling in this area of those unique Mac address.

We extract the most frequency detection Mac address. There are nine Mac address were detected for 102 days among all five months of data collection period. We track those Mac address to obtain their OD trip pattern behavior and estimate their route choice behaviors. Figure 19 demonstrate OD pattern from 1st Mac address to 9th Mac address. These OD patterns show whether where those Mac address usually travel from which area to which area. Those Mac addresses of 1st, 2nd, 3rd, 4th, 5th and 8th present several of unconformities trip patterns while the others 6th, 7th and 9th generate quite conformation patterns. Those six Mac addresses generate from many origins to many destinations, travel with no usual patterns. However, it is clearly seen that those remaining three Mac address especially the 9th Mac address depart from device ID 211 and reach to the adjacent device ID 248 mostly.

![Figure 18: Frequency unique Mac address count with number of detection days](image-url)
Figure 19: OD pattern of unique MAC address at the most frequency detection at 102 days
We demonstrate route choice behavior on the map of the maximum OD trips for the 1st Mac address between OD pair 248-252, figure 20. Line and number represent tracking route of visited links and their direction sequences. There are five trips generated on this OD pair as it is the largest OD trips in this Mac address represented on red line in figure 15. The 1\textsuperscript{st} and 2\textsuperscript{nd} OD trip travel simply pattern from 248-247-106-252 and 248-247-106-252 respectively. The 3\textsuperscript{rd} OD trip is most complicated travel visits, total 18 visited links. The trip head to South and cross the bridge to West. After that visited duplicate a few times along 238, 232 and 253 then head to the destination 252. The 4\textsuperscript{th} and 5\textsuperscript{th} trip present similarly route behavior pattern. They start the trip from origin head to South directional to 235. Then turn to the destination 252 as circle loop trip pattern.

![Route choice behavior of 1st Mac address, OD pair 248-252](image)

**Figure 20:** Route choice behavior of 1st Mac address, OD pair 248-252

7. Conclusion and discussion

As the proper traffic control needs to obtain the fundamental traffic data. However, the existing traffic data collection method has high introduction cost, and detailed traffic conditions cannot be update information under uncommon circumstance such as closing road or accidents. Many cities in Southeast Asia have not had adequate vehicle sensors lead to inappropriate signal control. Bangkok,
Thailand is well known as the top among most traffic congestion city in the world and no proper traffic detection system. Therefore, this study would like to overcome above mentioned problem by implement the Bluetooth technology which represents rather less cost than the others in Bangkok as a case study for long period. We installed several Bluetooth detectors to cover important intersection on the arterial road network. Travel time was extracted from raw data recorded as time stamp at each device through the server.

The raw summary data of detector operation and capability detection amount are used for screening outlier in primary. The trip between origins to destinations was defined as complete detection and missing link detection. We completely solved these missing links by shortest path algorithm beforehand. Then, we calculated speed using those shortest distance estimations divided by travel time from our Bluetooth time stamp detection. The estimated speed is used for screening outlier base on real experimental assumption from previous study.

We completely able to extract the traffic data which are OD pattern, speed pattern and tracking individual Mac address OD pattern. Those traffic data can be represented for the pattern of actual traffic data. The Bluetooth data collection can help the traffic planner to grasp the traffic situation updated time by time unlike doing questionnaire surveys. Some special holidays lead to less detection trending that can be seen from Bluetooth data collection but will not see from taking questionnaire. The OD pattern can guide the traffic planner to define the most trending possible people travel from where and to where. Therefore, the traffic control can be improved properly according to refer those OD patterns.

The speed pattern able to present the overview traffic speed that can be used for identification capability flow in this study area. Long term period Bluetooth data collection performed usefulness of individual unique Mac address OD pattern and route choice behaviors. In contrast with GPS data tracking that commonly install in business firm, this data from Bluetooth collection can help the traffic planner to track the route choice behavior of passenger driver which is valid under the privacy issues. Further, these route choice behaviors can be analyzed more to see the traffic data pattern by each link.

Unfortunately, we could not investigate how proportion of people turn on Bluetooth signal while travelling. Therefore, we could not exactly observe ability of detection rate. The increasing adaption of this technology possibly carries more detection rate and lead to better accuracy data. However, the unidirectional antenna, which was indicated has significantly higher fraction read than any other antenna, is installed to this study to catch huge detections. The suggestion to be able to increase potential of detection rate during heavy traffic stream is to install multiple Bluetooth readers in one location.

ACKNOWLEDGEMENTS

This study was carried out with the fund from the Japan Society of Traffic Engineer. We would like to thank the JSCE. Also, we would like to express our gratitude to Assoc. Prof. Sorawit Narupiti, Mr. Onnz Kittipong and the students of Chulalongkorn University for their cooperation to carry out data collection and analysis.

References


Origin-destination trip ... 12 (2021) No. 2, 1541–1562


Appendix

Raw data

Bluetooth data collection in 2018 is the collaborative research between Chulalongkorn University (CU), Nihon University (NU) and other universities. Our aim is to explain the current situations for collecting Bluetooth data in downtown of Bangkok, Thailand in 2018. The objectives of this project are (1) to develop the equipment monitoring Bluetooth devices in long term approximately six months and (2) to develop the methodology for estimating travel time, origin destination matrix and forecasting route choice.

This study begins to employ installation and testing the new system of Bluetooth detectors along the main road networks in Bangkok from October 2017 and completely installed all detectors in December 2017 for totally 45 detectors. However, Some Bluetooth detectors in previous remain...
disable due to problems of internet and signal device. Therefore, the new system of equipment is installed. The Bluetooth detectors in this phase are embed 4G communication module to the raspberry pi compute board and developed in term of ship processing to be own Bluetooth scanning device as shown in figure 1-1. This study expands the location of important intersections along road networks and put all detectors in the police box, the name of location of Bluetooth scanner are as shown in Table 1.

The primary test of Bluetooth data is reported from database starting from December 29, 2017 to January 13, 2018 totally 16 days. Figure 1-3 shows the example of raw data gotten from database server. Regarding the Bluetooth data, Tohoku university kindly provides the completed filtering and cleansing data as well as split trips depending on unique MAC address and date in.

The number of raw data and unique MAC address data that each scanner can collect during the data collection period can be presented in table 2 and table 3. The finding shows that the trend of data varies between weekday and weekend indicating that the number of traffic data in weekend (i.e. 6, 7 and 13) are less than its weekday. The facts also show that although the day on January 1 and 2, 2018 are weekdays, both remain in the New Year holiday that people normally go back to their hometown outside the Metropolitan.

Regarding the capability of Bluetooth detectors, they can detect Bluetooth devices from travelers on the average 20,000 unique Mac address per day. This study should get the actual traffic volume in each intersection in order to determine the penetration rate and define the capability of Bluetooth detectors.
Table 1: List of Bluetooth detector location

<table>
<thead>
<tr>
<th>Detector</th>
<th>Intersection</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>Henry Dunant</td>
</tr>
<tr>
<td>106</td>
<td>Chaloem Pao</td>
</tr>
<tr>
<td>108</td>
<td>Samyan</td>
</tr>
<tr>
<td>109</td>
<td>Saphan Laueng</td>
</tr>
<tr>
<td>111</td>
<td>Maha Nakorn</td>
</tr>
<tr>
<td>201</td>
<td>Rachprasong</td>
</tr>
<tr>
<td>202</td>
<td>Phraram 4</td>
</tr>
<tr>
<td>203</td>
<td>Langsuan</td>
</tr>
<tr>
<td>205</td>
<td>Nana</td>
</tr>
<tr>
<td>207</td>
<td>Soi Sukhumvit42</td>
</tr>
<tr>
<td>210</td>
<td>Asok-Phraram</td>
</tr>
<tr>
<td>211</td>
<td>Phong Phraram</td>
</tr>
<tr>
<td>212</td>
<td>Makkasan</td>
</tr>
<tr>
<td>214</td>
<td>Tai Duan Phetburi</td>
</tr>
<tr>
<td>215</td>
<td>Phrayathai</td>
</tr>
<tr>
<td>216</td>
<td>Phatunam</td>
</tr>
<tr>
<td>217</td>
<td>Yommarach</td>
</tr>
<tr>
<td>218</td>
<td>Pleojit</td>
</tr>
<tr>
<td>219</td>
<td>Mitsumphan</td>
</tr>
<tr>
<td>221</td>
<td>Kasatsruk</td>
</tr>
<tr>
<td>222</td>
<td>Hua Lumpong</td>
</tr>
<tr>
<td>225</td>
<td>Surasak</td>
</tr>
<tr>
<td>226</td>
<td>Mahasak</td>
</tr>
</tbody>
</table>

Figure 1-3
### Table 2: Number of traffic data in 2018

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw data</td>
<td>102,810</td>
<td>98,594</td>
<td>94,296</td>
<td>89,750</td>
<td>80,775</td>
<td>98,611</td>
<td>103,825</td>
<td>111,284</td>
<td></td>
</tr>
<tr>
<td>Unique MAC address</td>
<td>24,002</td>
<td>21,935</td>
<td>20,058</td>
<td>19,293</td>
<td>18,538</td>
<td>23,527</td>
<td>24,662</td>
<td>26,749</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3: Number of traffic data in 2018 (cont)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw data</td>
<td>198,206</td>
<td>91,832</td>
<td>99,659</td>
<td>97,180</td>
<td>95,711</td>
<td>101,572</td>
<td>101,837</td>
<td>84,994</td>
</tr>
<tr>
<td>Unique MAC address</td>
<td>23,363</td>
<td>21,735</td>
<td>24,319</td>
<td>24,222</td>
<td>24,151</td>
<td>25,623</td>
<td>25,640</td>
<td>21,294</td>
</tr>
</tbody>
</table>