Int. J. Nonlinear Anal. Appl. 12 (2021) No. 2, 1541–1562
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/IJNAA.2021.5282

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

Rattanaporn Kasemsri[a,*], Tetsuhiro Ishizaka[b], Atsushi Fukuda[b], Piyapong Suwanno[c]

[a] *School of Civil Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima, 30000, Thailand.*
[b] *College of Science and Technology, NihonUniversity, Chiba, 274-8501, Japan.*
[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

*Corresponding Author: Rattanaporn Kasemsri
Email addresses: rattanaporn7777@gmail.com (Rattanaporn Kasemsri[a,*]), ishizaka.tetsuhiro@nihon-u.ac.jp (Tetsuhiro Ishizaka[b]), fukuda.atsushi@nihon-u.ac.jp (Atsushi Fukuda[b]), piyapong.su@rmutsv.ac.th (Piyapong Suwanno[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 Bluetooh 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).

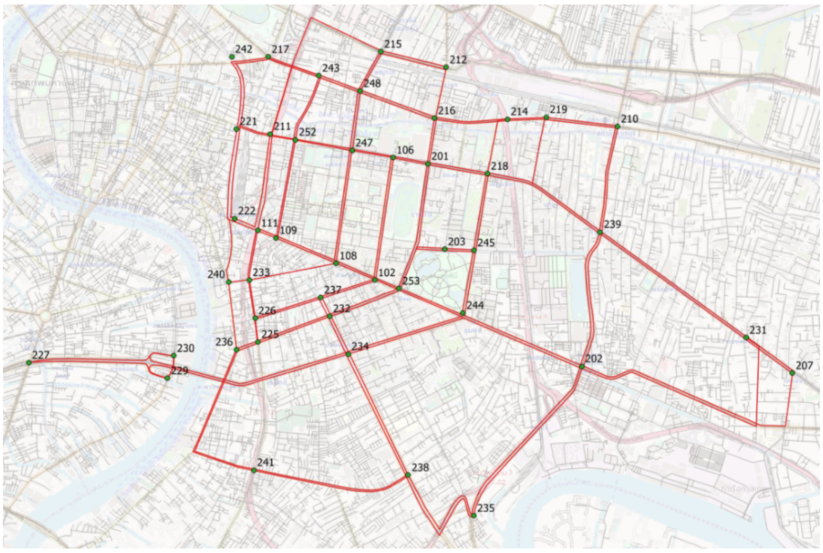


Figure 1: Bluetooth detector locations in Bangkok, Thailand

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.

	V1	V2	V3	V4	V5	V6	V7
1	153153122	-	94:D0:29:38:D8:0E	-70	2018-01-14 00:00:00.758222	2018-01-14 00:00:08.000000	215
2	153153123	-	94:D0:29:38:D8:0E	-55	2018-01-14 00:00:03.909322	2018-01-14 00:00:08.000000	215
3	153153124	-	94:D0:29:38:D8:0E	-68	2018-01-14 00:00:04.841272	2018-01-14 00:00:08.000000	215
4	153153125	-	94:D0:29:38:D8:0E	-66	2018-01-14 00:00:05.161221	2018-01-14 00:00:08.000000	215
5	153153126	-	94:D0:29:38:D8:0E	-64	2018-01-14 00:00:05.880324	2018-01-14 00:00:08.000000	215
6	153153127	-	94:D0:29:38:D8:0E	-54	2018-01-14 00:00:07.160291	2018-01-14 00:00:08.000000	215

Figure 2: Example of raw data from database server

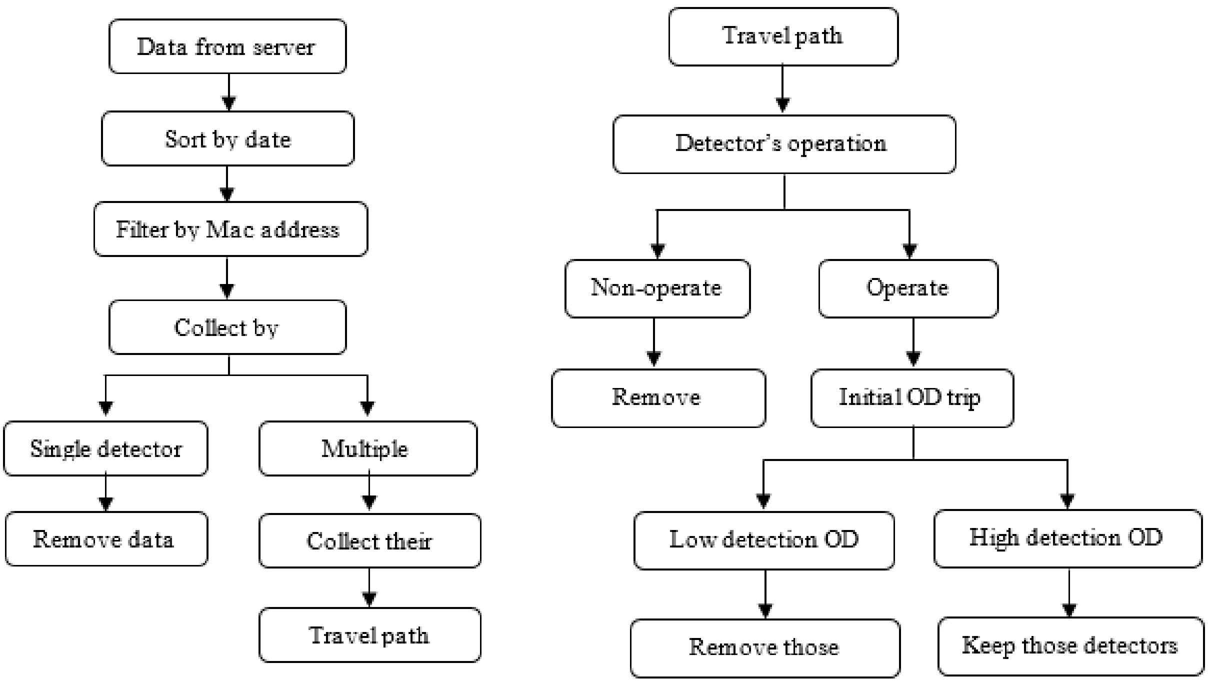


Figure 3: Extraction data process from database server

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.

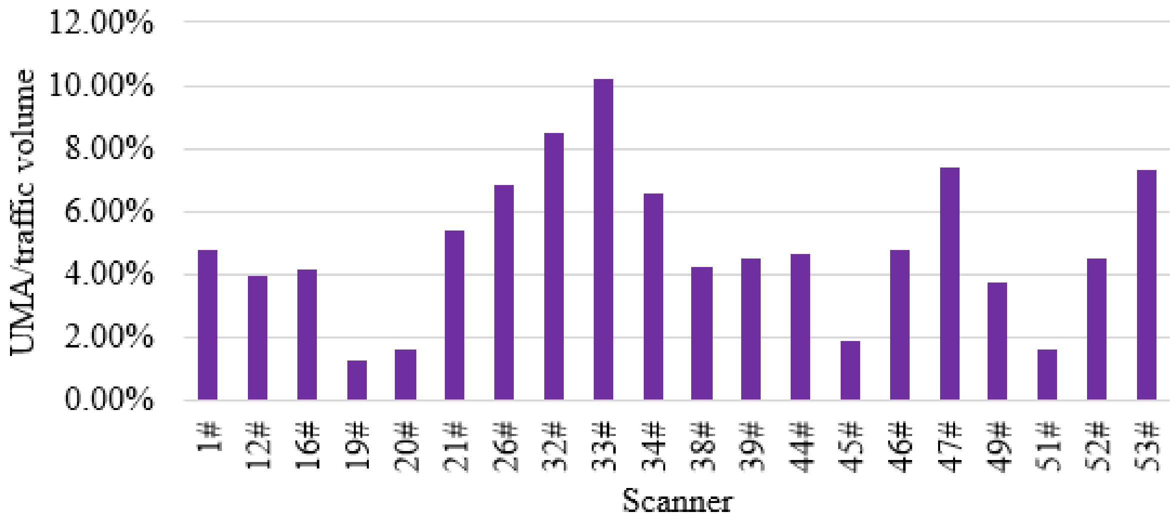


Figure 4: Traffic volume V.S. Unique Mac address

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

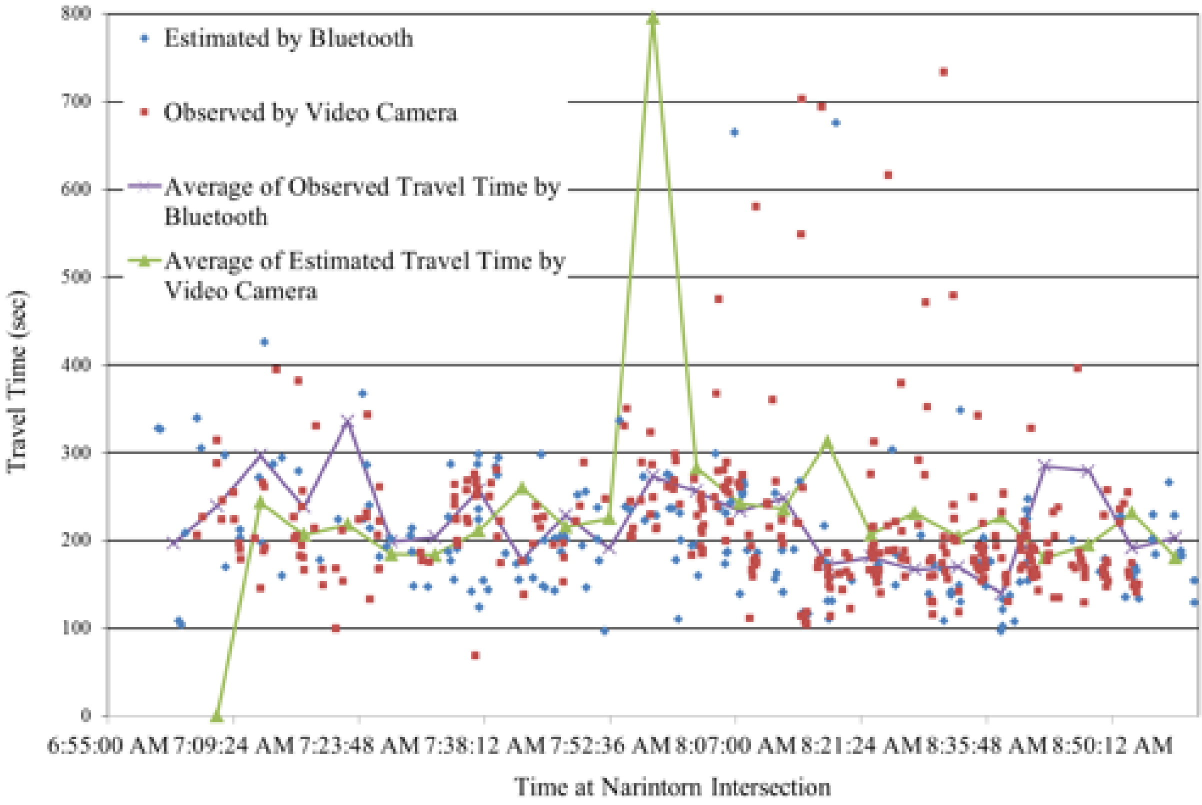


Figure 5: Estimated travel time by bluetooth and video camera

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.

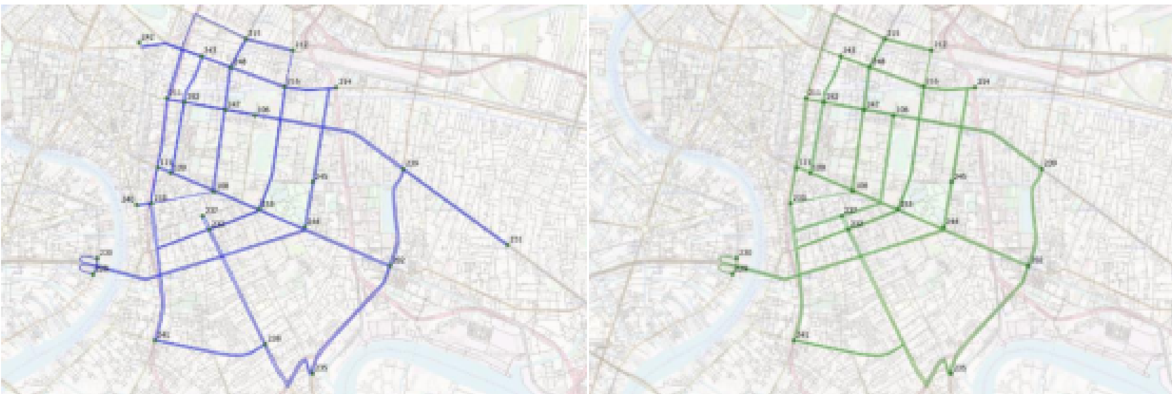
5.1. Detector's operation

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 7b.

Origin \ Destination	106	108	109	111	202	211	212	213	214	215	216	218	229	230	231	232	233	235	236	237	238	239	240	241	242	243	244	245	246	247	248	252	253
106	NA	8376	2576	2203	3146	1290	8217	15	5863	657	803	68	817	192	1256	2087	1179	1094	11	1147	1071	1485	2097	1169	5576	971	2129	3272	NA	5821	4258	8564	5239
108	7841	NA	16653	15129	8330	779	5247	7	3494	558	170	36	2396	451	1268	6845	6640	2994	22	4761	3806	782	3515	3900	3760	503	7122	4294	NA	4188	3942	6493	8964
109	1131	7787	NA	8207	2408	351	891	9	559	177	43	7	568	116	372	744	2871	1060	24	564	716	220	429	1612	797	184	1731	941	NA	599	614	1788	1099
111	1728	10943	9278	NA	3241	998	2029	2	1002	319	172	40	796	335	547	1530	5874	1245	33	1314	1211	437	711	1544	2062	152	2689	1218	NA	1054	1241	2051	2336
202	2426	9654	3263	4921	NA	587	5888	3	10186	387	375	132	3575	1001	13425	5397	3718	18792	45	2019	6687	9290	4293	5077	3832	558	18566	10415	NA	1841	2154	3697	13212
211	330	249	236	461	170	NA	529	2	184	74	70	12	102	56	77	97	228	101	5	86	73	89	48	134	1387	67	147	54	NA	658	337	1650	129
212	4035	2404	715	1295	5118	1415	NA	14	17681	1998	6781	286	1385	435	2769	2639	1662	2683	40	1124	1528	6528	4801	2446	6947	1358	3802	7402	NA	3497	10211	4906	3750
213	1	1	1	NA	NA	NA	15	NA	NA	NA	1	NA	NA	NA	NA	1	3	1	NA	NA	1	NA	NA	NA	14	1	3	1	NA	NA	8	2	NA
214	6087	3713	1192	933	11908	616	20062	NA	NA	1241	2088	213	1786	423	5447	3047	1894	5968	41	1222	2507	8184	4009	2955	10353	3988	5898	18138	NA	2748	8145	4994	6555
215	1937	2010	702	743	1339	879	11591	NA	5539	NA	676	26	425	166	777	807	816	619	38	453	558	674	638	797	1973	1580	1032	686	NA	3821	7537	3536	1021
216	398	221	87	173	337	267	6428	5	1136	104	NA	61	125	54	297	227	115	149	3	79	139	518	266	182	1687	200	290	307	NA	417	2917	587	319
218	120	45	16	37	157	31	492	NA	279	14	76	NA	19	9	97	38	20	45	2	21	48	262	51	38	90	14	142	231	NA	82	94	65	144
229	820	2497	950	764	3851	275	2050	NA	2771	135	131	15	NA	36595	1018	4000	4335	5543	94	1821	3368	603	1190	7628	2122	220	3656	2720	NA	651	864	1100	1596
230	250	616	233	516	1305	204	588	NA	633	85	65	12	51911	NA	377	1018	931	1671	29	503	1061	234	283	2624	1088	42	1083	577	NA	272	299	432	484
231	2524	2296	821	1149	25660	456	5109	5	10329	381	676	280	1384	459	NA	1823	1174	7764	23	736	2668	10295	1403	2408	2902	521	4626	5028	NA	1626	2136	2003	3521
232	2526	7730	1502	1330	5545	428	5099	NA	3960	327	374	44	4166	808	1991	NA	5634	2812	119	17406	11661	1026	5957	7359	2178	294	7714	3799	NA	1564	1777	2243	10766
233	2122	17092	6379	5917	6159	906	3880	2	3762	536	174	28	6836	1259	1053	13854	NA	3023	294	10566	6952	758	2213	17272	4414	548	3994	3318	NA	2642	2668	5226	3809
235	2053	6726	2814	1778	49500	354	5098	3	9218	286	296	96	5526	1222	10288	3932	2478	NA	29	1607	20929	6323	3814	10934	2998	371	8422	12484	NA	1154	1684	2378	8618
236	11	56	65	35	70	14	51	NA	74	19	5	NA	95	33	21	126	317	38	NA	57	89	18	36	262	6	22	40	19	NA	26	23	25	74
237	1161	4574	655	551	1389	116	1068	1	948	98	67	12	1169	223	296	6726	3434	647	18	NA	2525	237	883	2041	707	129	1335	906	NA	687	692	923	2106
238	1006	2829	1028	800	8172	176	2149	1	2632	160	246	47	3294	1085	2012	9075	2472	12936	38	4136	NA	922	1621	17984	1048	145	7126	3889	NA	613	930	889	3845
239	3105	2042	727	957	22205	776	9751	2	21350	397	973	825	1076	393	9059	1897	999	4109	27	895	2370	NA	1730	1873	3383	475	7024	10268	NA	2274	2678	3167	3581
240	1849	3094	877	838	4659	206	7576	2	4913	168	479	115	1060	247	920	4842	1075	1524	35	1130	1810	1516	NA	1487	1349	302	4865	10772	NA	904	1362	1716	11179
241	1477	5030	2549	1147	7538	255	3494	3	4230	334	263	41	11022	2656	2130	7346	12669	11780	304	3789	20682	1113	1993	NA	3000	452	5172	4503	NA	1019	1586	1697	2986
242	1396	1944	1102	1564	3560	1786	11094	19	8497	590	744	45	3257	741	1600	1010	3092	2259	5	608	989	1319	782	3337	NA	1846	1578	1282	NA	2006	5270	4747	1096
243	2948	2216	1150	930	2397	1169	9686	NA	16438	2101	724	36	823	245	1074	1414	1526	1263	55	698	700	1369	889	1276	9826	NA	1796	2127	NA	2602	8146	9545	1297
244	2062	5681	2025	3177	31852	615	6367	3	14611	322	634	287	4278	1272	4291	7191	1975	7421	37	2052	9266	3439	5513	5581	3285	348	NA	30157	NA	1513	2063	2634	11820
245	2395	3732	1325	1091	12257	186	6399	4	22742	176	244	179	2522	471	2246	3458	1767	4929	18	1050	4584	4041	6448	3656	2458	458	21395	NA	NA	1264	1434	2560	5398
246	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
247	2657	2432	868	885	1461	1785	4407	6	1763	825	305	25	531	114	599	805	1078	512	9	447	427	760	617	633	4381	589	830	945	NA	NA	5562	12399	827
248	7743	9370	1916	2644	4812	3358	33111	66	10307	2173	2754	101	1792	460	2323	2992	2778	2120	12	1856	2099	2129	2870	2866	15869	2949	3765	3453	NA	15269	NA	14619	3590
252	3585	5236	2600	2497	3026	5297	5946	26	3496	548	305	21	1180	266	679	1672	4000	1228	15	1175	1058	703	956	1906	10604	1834	2206	1511	NA	7096	4426	NA	1522
253	3973	8229	2317	2183	13756	351	6851	1	6246	283	524	71	2124	429	2039	12100	1999	4209	42	3882	5645	1292	14139	3305	1865	282	15202	7096	NA	1481	1786	2174	NA

Figure 6: 6OD table receiving from primary data



a Detector's operation b Unconnected link

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 14^{th} and 15^{th}, which is correspondingly with national holidays.

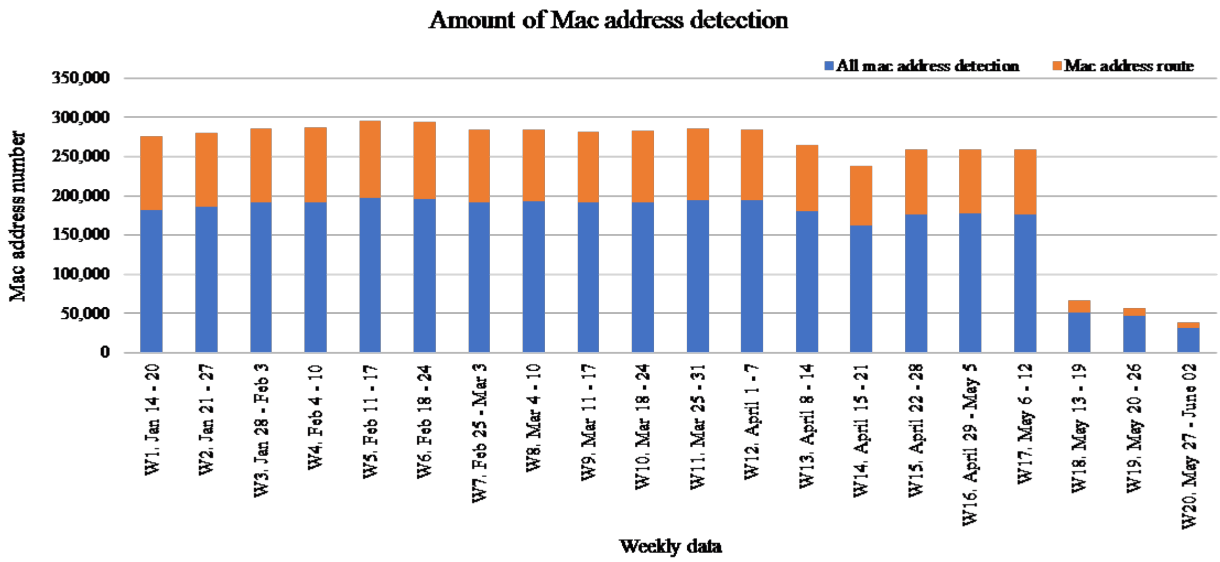


Figure 8: The proportion amount of detection by single detection and multiple detections

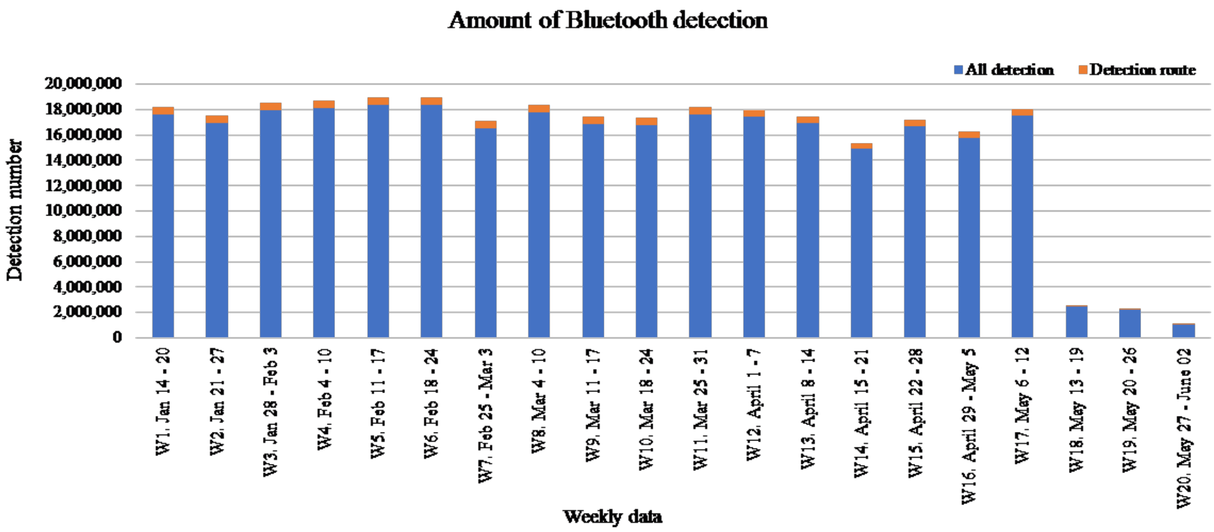


Figure 9: The proportion of unique Mac address by single detection and multiple detections

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.

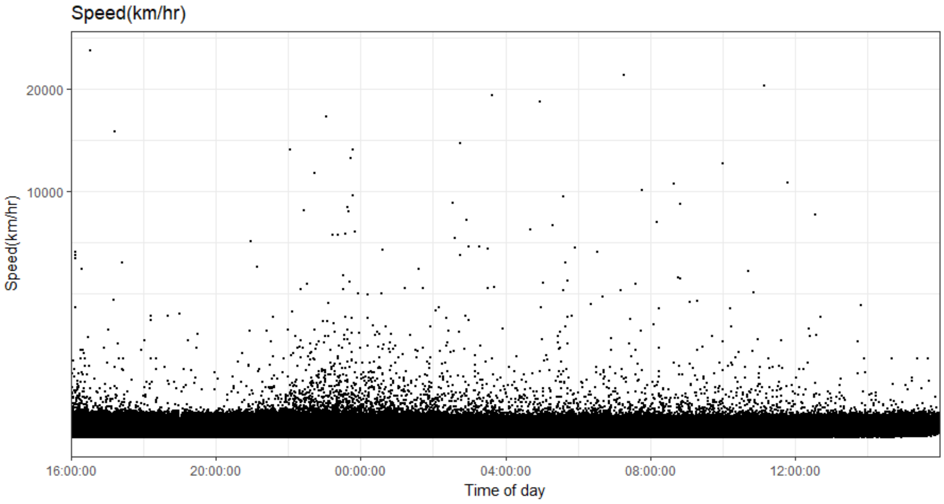


Figure 10: Calculation speed from shortest path estimation

Absolutely seen at above of thousand $km/hour$ 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 $km/hour$. 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.

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

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 pairon trip k.

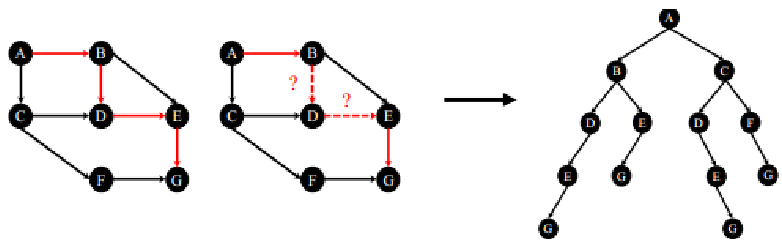


Figure 11: An example of the missing link and BFS traverses a graph

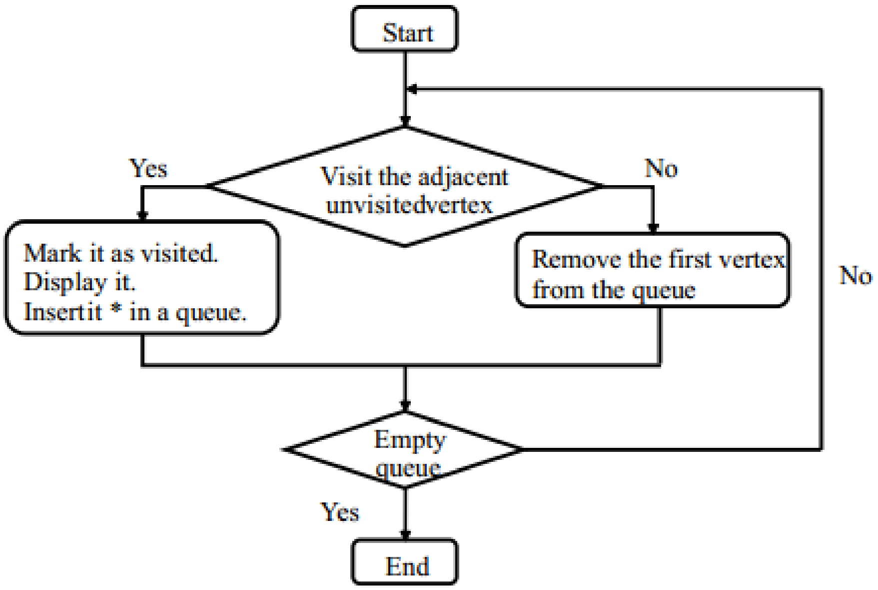


Figure 12: The rules of BFS algorithm

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 breadth ward 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 to 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.

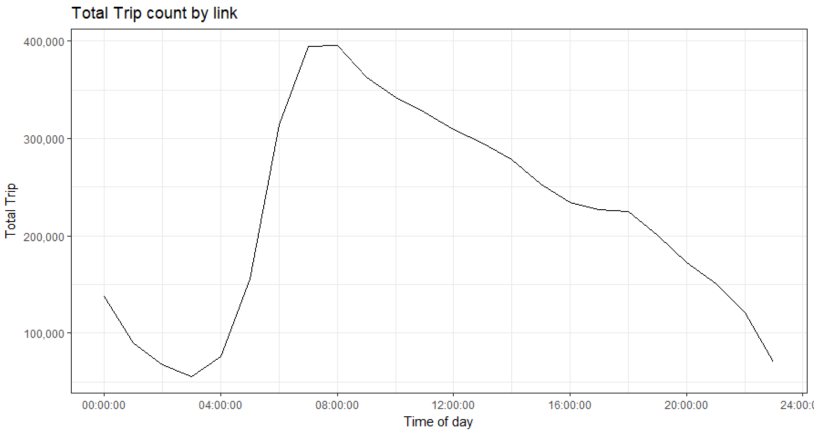


Figure 13: The number of link count trips by time of day

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 14demonstrate for OD pattern of every weekend.

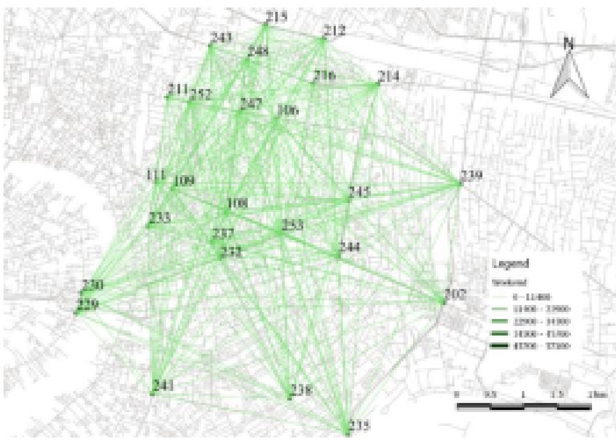


Figure 14: Weekend OD patterns

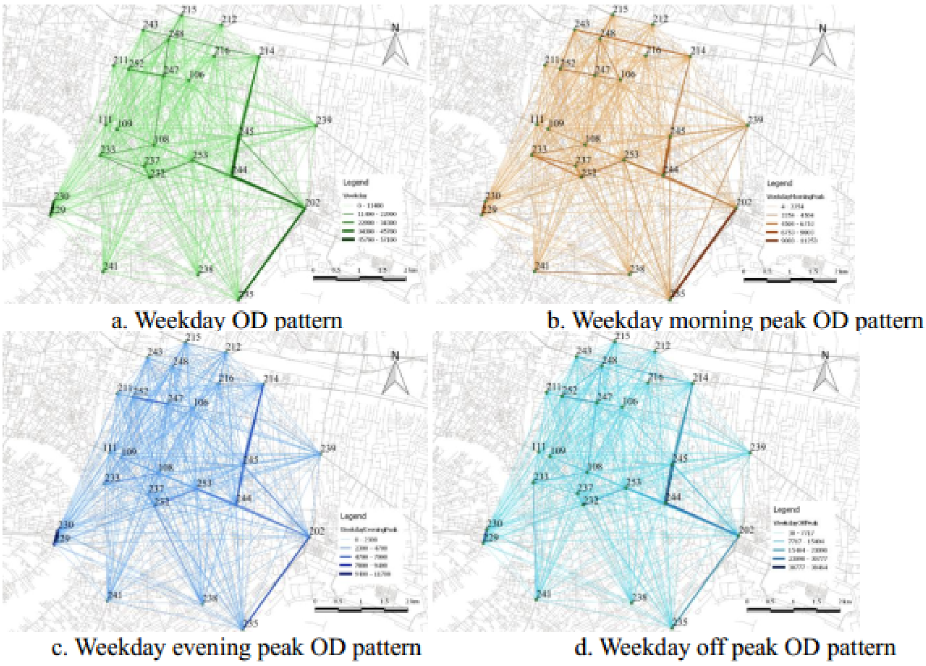


Figure 15: Weekday OD patterns

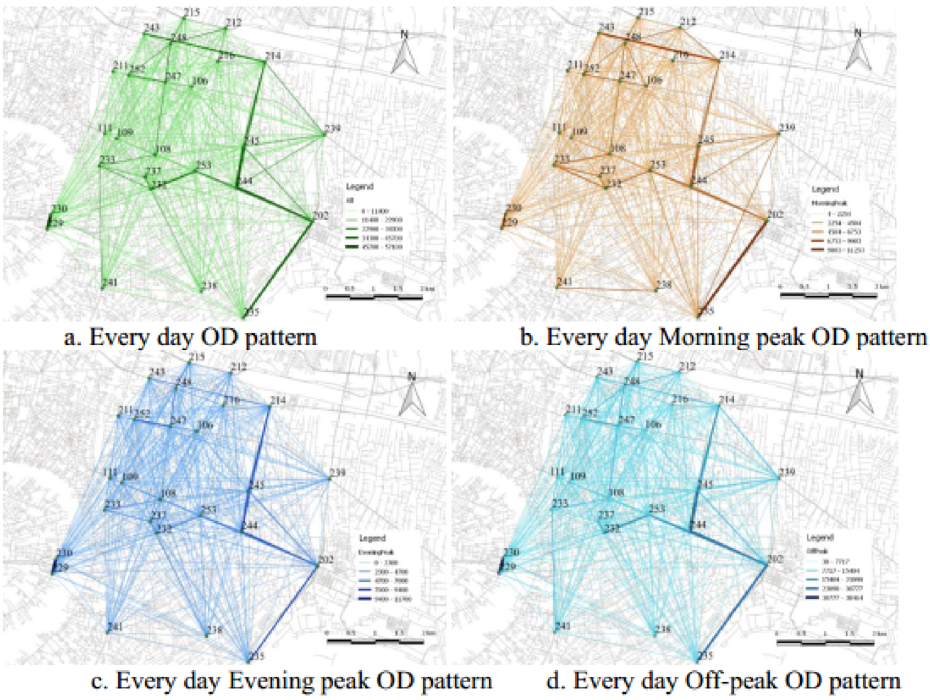


Figure 16: OD patterns of every day

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.

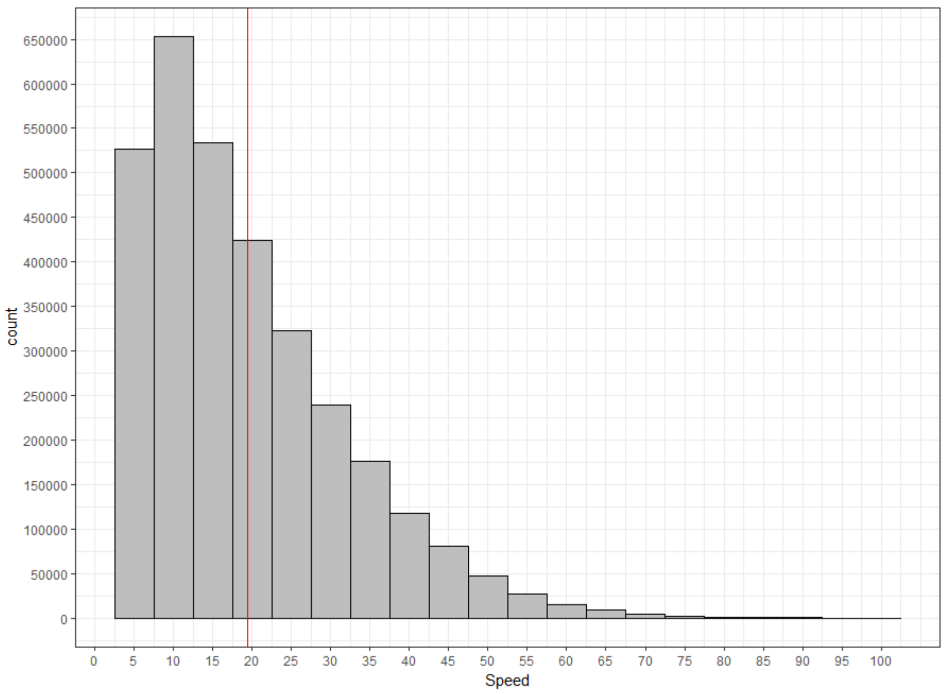


Figure 17: Frequency driving speed and average speed

6.3. Mac address frequency

Specific information is the frequency classified by unique Mac address as shown inFigure 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 1^{st} Mac address to 9^{th} Mac address. These OD patterns show whether where those Mac address usually travel from which area to which area. Those Mac addresses of 1^{st}, 2^{nd}, 3^{rd}, 4^{th}, 5^{th} and 8^{th} present several of unconformities trip patterns while the others 6^{th}, 7^{th} and 9^{th} 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 9^{th} Mac address depart from device ID 211 and reach to the adjacent device ID 248 mostly.

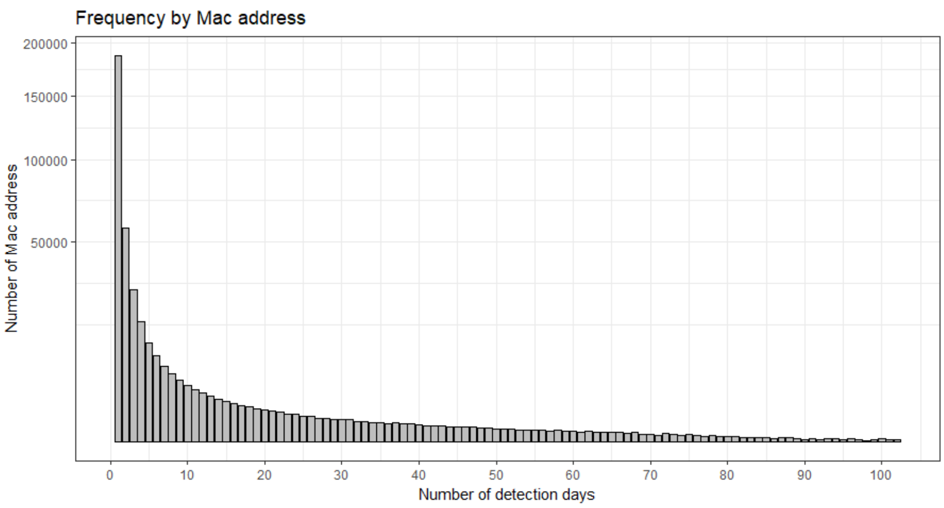


Figure 18: Frequency unique Mac address count with number of detection days

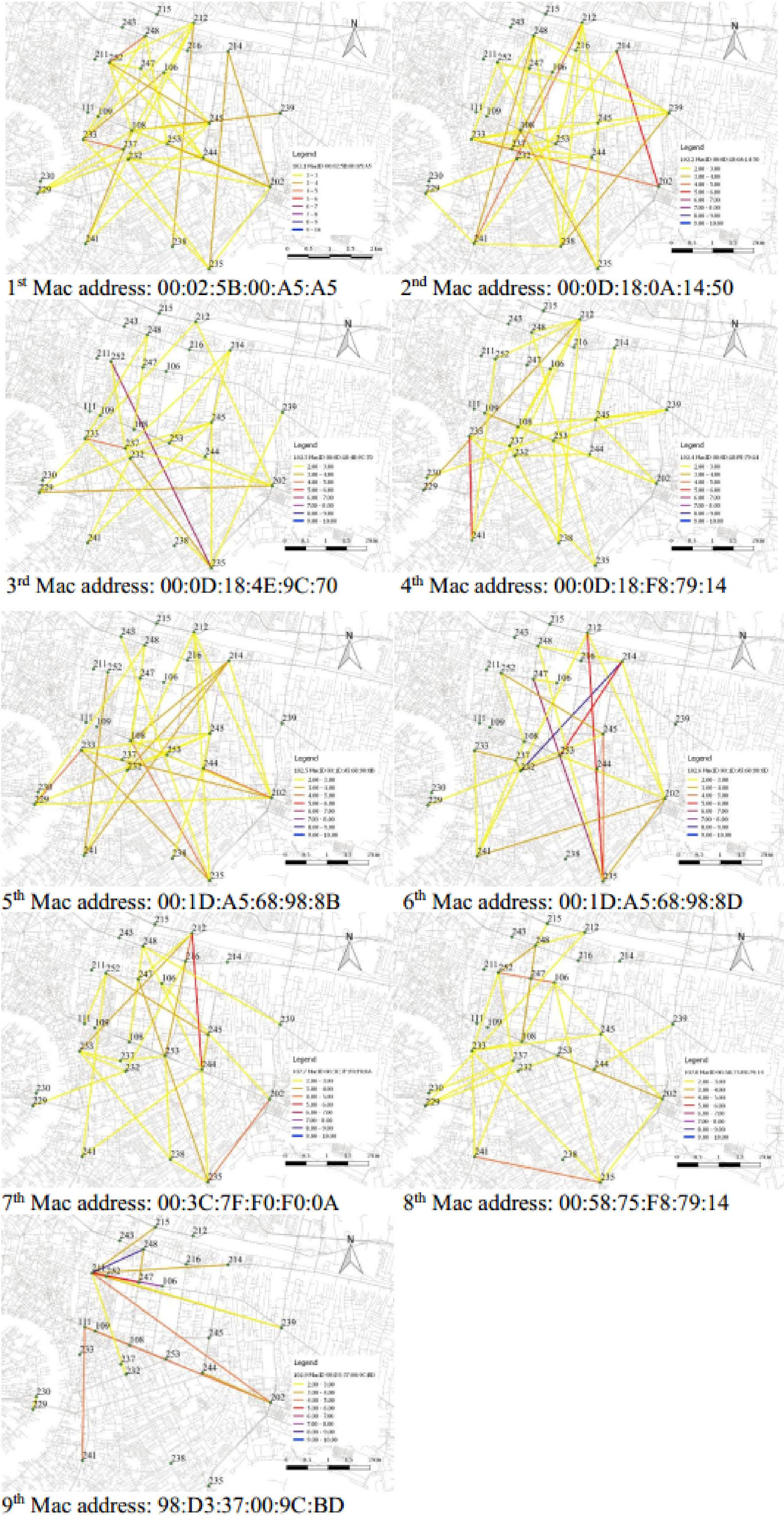


Figure 19: OD pattern of unique Mac address at the most frequency detection at 102 days

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^{st} and 2^{nd} OD trip travel simply pattern from 248-247-106-252 and 248-247-106-252 respectively. The 3^{rd} OD trip is most complicated travel visits, total 18 visited links. The trip head to Southand 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^{th} and 5^{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.

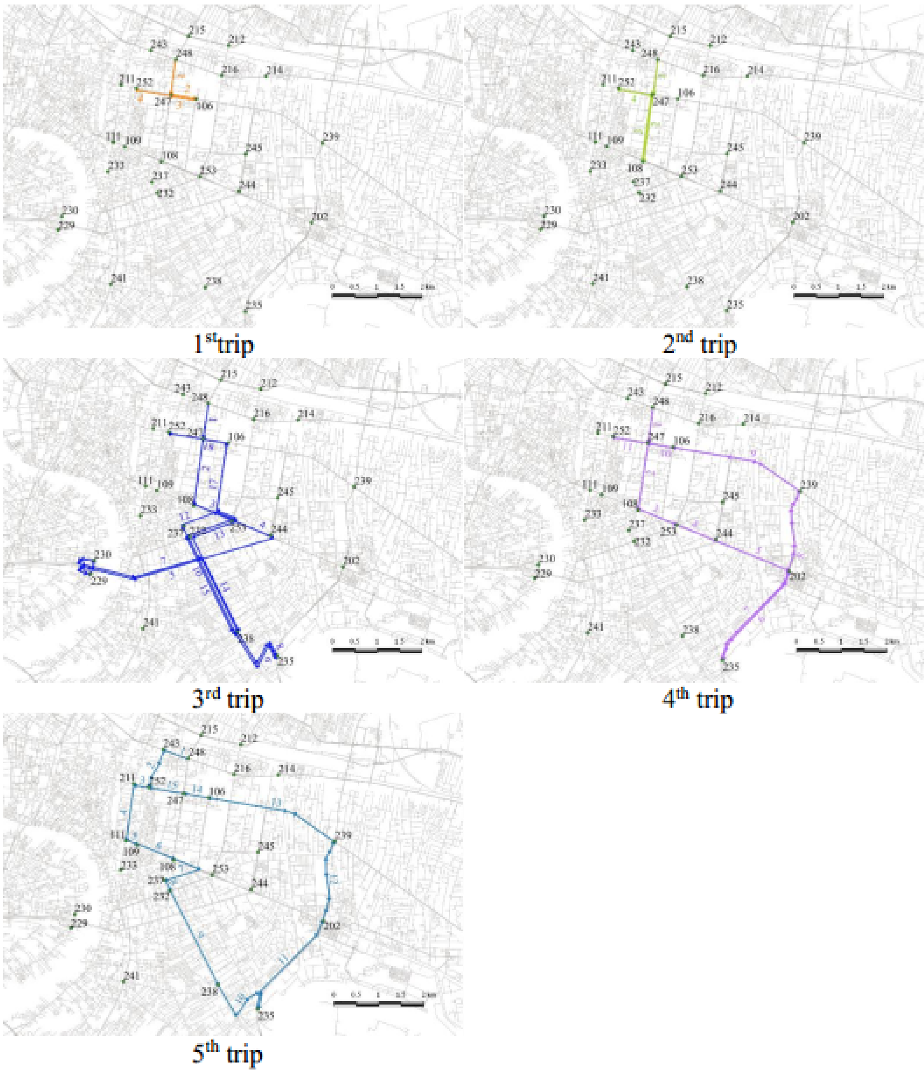


Figure 20: Route choice behavior of 1st Mac address, OD pair 248-252

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

[1] J.D. Porter, D.S. Kim, M.E Magana, P. Poocharoen, C. Antar and G. Arriaga, *Antenna Characterization for Bluetooth-based Travel Time Data Collection*,Journal of Intelligent Transportation Systems, 17 (2013) 142-151.

[2] B. N. Araghi, L. T. Christensen, R. Krishnan and H. Lahrmann, *Application of Bluetooth technology for mode-specific travel time estimation on arterial roads: Potentials and challenges*, In Proceedings from the Annual Transport Conference at Aalborg University, (2012).

[3] B.N. Araghi, K. S. Pedersen, L. T. Christensen, R. Krishnan, H. Lahrmann, *Accuracy of travel time estimation using Bluetooth technology: Case study Limfjord tunnel Aalborg*, International Journal of Intelligent Transportation Systems Research, 13(3) (2015) 166-191.
[4] W. Charle, F. Viti and C. Tampère, *Estimating route travel time variability from link data by means of clustering, Paper presented at 12th WCTR*, Lisbon, Portugal, July (2010) 11-15.
[5] Y. Zhang and A. Haghani, *A gradient boosting method to improve travel time prediction*, Transportation Research Part C: Emerging Technologies, 58 (2015) 308-324.
[6] L. Vanajakshi, S.C. Subramanian and R. Sivanandan, *Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses*, IET Intelligent Transport Systems, 3(1) (2009) 1-9.
[7] W.H. Schneider, S.M. Turner, J. Roth and J. Wikander, *Statistical Validation of Speeds and Travel TimesProvided by a Data Service Vendor*, The University of Akron, (2010).
[8] P. Slone, *Bluetooth Based Travel Time Analysis.Southern District ITEAnnual Meeting*, Lafayette, LA, (2011).
[9] S. Young, *Bluetooth Traffic Monitoring Technology - Concept of Operation &Deployment Guidelines*, University of Maryland - Center for AdvancedTransportation Technology, (2008).
[10] D.D. Puckett and M.J. Vickich, *Bluetooth-Based Travel Time/Speed MeasuringSystems Development. University Transportation Center for Mobility*, TexasTransportation Institute, (2010).
[11] KMJ. Consulting. *Bluetooth Travel Time Technology Evaluation-Using theBlueTOAD*, KMJ Consulting, Inc., Haverford, PA, (2010).
[12] E. Sharifi, M. Hamedi, A. Haghani and H. Sadrsadat, *Analysis of vehicle detection rate for Bluetooth traffic sensors: A case study in Maryland and Delaware*, In: 18thWorld Congress on Intelligent Transport Systems, (2011).
[13] M. Martchouk and F. Mannering, *Analysis of Travel Time Reliability on IndianaInterstates*, Purdue University, (2009).
[14] P.J. Tarnoff, D.M. Bullock, S.E. Young, J. Wasson, N. Ganig and J.R. Sturdevant, *Continuing Evolution of Travel Time Data InformationCollection and Processing*, Transportation Research Board 88th AnnualMeeting, Washington DC,14 (2009).
[15] J. Barceló, L. Montero, L. Marqués, P. Marinelli and C. Carmona, *Travel Time Forecasting and Dynamic OD Estimation inFreeways Based on Bluetooth Traffic Monitoring*, Department of Statistics andOperations Research and CENIT2009, (2009).
[16] Y. Malinovskiy, U.K. Lee, Y. Wu and Y. Wang, *Investigation of Bluetooth-Based Travel Time EstimationError on a Short Corridor*, Paper presented at the 90th Annual Transportation ResearchBoard Meeting, Washington, D.C., (2011).
[17] T.M. Brennan, M.E. Joseph, M.D. Christopher, M.B. Darcy , V.K. James and M. Mary , *Influence of Vertical Sensor Placement on Data CollectionEfficiency from Bluetooth MAC Address Collection Devices*, Journal ofTransportation Engineering, 136 (2010) 1104-1109.
[18] Y. Malinovskiy, Y. Wu, Y. Wang and U.K. Lee, *Field Experiments on Bluetooth-based Travel Time DataCollection*, Paper presented at the 89th Annual Transportation Research BoardMeeting, Washington, D.C., (2010).
[19] Y. Wang, Y. Malinovskiy, Y. Wu, U.K. Lee, *Field Experiments with Bluetooth Sensors*, (2010).
[20] *The Bluetooth Special Interest Group*, Architecture & Terminology Overview.Specification of the Bluetooth System,1 (2004) 13-14.

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 projectare (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 moduleto the raspberry pi compute board and developed in term of ship processing to beown Bluetooth scanning device as shown in figure 1-1. This study expands thelocation of important intersections along road networks and put all detectors in thepolice box, the name oflocation of Bluetooth scannerare as shown in Table 1.

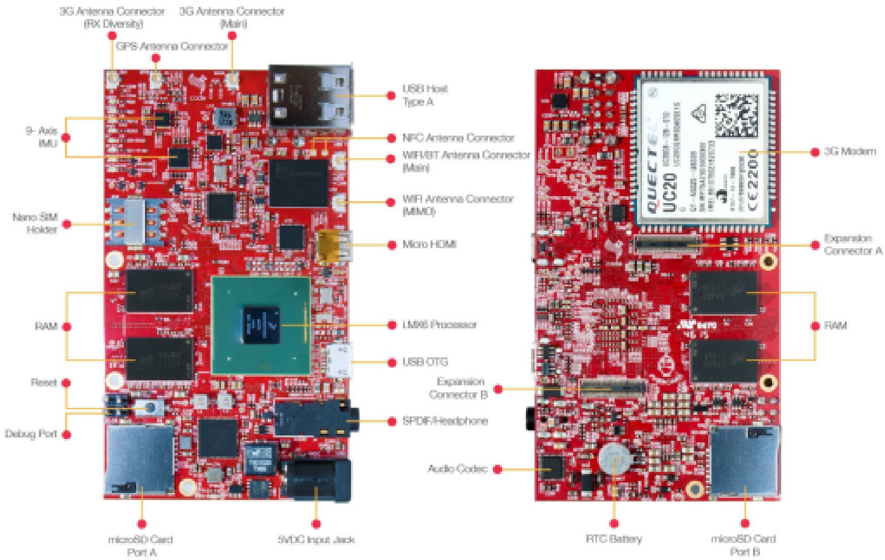


Figure 1-1 Ship processing in Bluetooth box

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

Detector	Intersection	Detector	Intersection
102	Henry Dunant	227	Taksin
106	Chaloem Pao	229	Klongtonsai
108	Samyan	230	Chareon Nakorn
109	Saphan Laueng	231	Thong Lor
111	Maha Nakorn	232	Silom-Narathiwas
201	Rachprasong	233	Maha Nakorn-Si Phraya
202	Phraram 4	234	Narinthorn
203	Langsuan	235	Nang Linchi Intersection
205	Nana	236	Bangrak
207	Soi Sukhumvit42	237	3Suravong-narathivas
210	Asok-Phraram	239	Chan-Narativas
211	Phong Phraram	239	Asoke
212	Makkasan	240	3Sarasin-Ratchadumri
214	Tai Duan Phetburi	241	Chan Road-Charoenrat intersection
215	Phrayathai	242	Saphan Khaw
216	Phatunam	243	Phetphraram
217	Yommarach	244	Witthayu
218	Pleojit	245	Sarasin
219	Mitsumphan	247	Pathumwan
221	Kasatsruk	248	Ratchathewee
222	Hua Lumpong	252	Chareon Phol
225	Surasak	253	Sala Daeng
226	Mahasak		

No.	Mac Address	Signal	Date	Server Received	Device ID
25180117	XX:5E:1C:97:3F:F1	-76 dbm	2017-11-23 0:44:36.880251	2017-11-23 0:45:34.000000	235
25180116	XX:5E:1C:97:3F:F1	-72 dbm	2017-11-23 0:44:36.750274	2017-11-23 0:45:34.000000	235
25180115	XX:5E:1C:97:3F:F1	-71 dbm	2017-11-23 0:44:36.509311	2017-11-23 0:45:34.000000	235
25180114	XX:5E:1C:97:3F:F1	-75 dbm	2017-11-23 0:44:36.198243	2017-11-23 0:45:34.000000	235
25180113	XX:9F:05:2F:A3:B6	-71 dbm	2017-11-23 0:44:36.188500	2017-11-23 0:45:34.000000	235
25180112	XX:5E:1C:97:3F:F1	-78 dbm	2017-11-23 0:44:35.597258	2017-11-23 0:45:34.000000	235
25180111	XX:9F:05:2F:A3:B6	-74 dbm	2017-11-23 0:44:34.979676	2017-11-23 0:45:34.000000	235
25180110	XX:9F:05:2F:A3:B6	-73 dbm	2017-11-23 0:44:30.876682	2017-11-23 0:45:34.000000	235

Figure 1-3

Table 2: Number of traffic data in 2018

Data	**29/12/2017**	**30/12/2017**	**31/12/2017**	**1/1/2018**	**2/1/2018**	**3/1/2018**	**4/1/2018**	**5/1/2018**
Raw data	102,810	98,594	94,296	89,750	80,775	98,611	103,825	111,284
Unique MAC address	24,002	21,935	20,058	19,293	18,538	23,527	24,662	26,749

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

Data	**6/1/2018**	**7/1/2018**	**8/1/2017**	**9/1/2018**	**10/1/2018**	**11/1/2018**	**12/1/2018**	**13/1/2018**
Raw data	198,206	91,832	99,659	97,180	95,711	101,572	101,837	84,994
Unique MAC address	23,363	21,735	24,319	24,222	24,151	25,623	25,640	21,294