# On The Characteristics of Wi-Fi Packet Sensors for Traffic and Pedestrian Flow Observation

Guanghui ZHOU[1], Fumitaka KURAUCHI[2] and Shuhei MYOKO[3]

[1]Member of JSCE, Dept. of Mechanical and Civil Eng., Gifu University
(1-1 Yanagido, Gifu City 501-1193, Japan)
E-mail: x3912005@edu.gifu-u.ac.jp
[2]Member of JSCE, Professor, Dept. of Civil Eng., Gifu University
(1-1 Yanagido, Gifu City 501-1193, Japan)
E-mail: kurauchi@gifu-u.ac.jp
[3]Member of JSCE, Dept. of Civil and Environmental Eng., Gifu University
(1-1 Yanagido, Gifu City 501-1193, Japan)
E-mail: x4523037@edu.gifu-u.ac.jp

To observe pedestrian and/or vehicle traffic volume, we often use automatic counting techniques. This study focuses on Wi-Fi packets sensor technology that records Wi-Fi probe requests from devices attracts researchers' attention. However, many of the research just use the sensor data directly without considering the difference of detection ability of sensors. In fact, the detection ability of sensors may change with the surrounding conditions. This research aims to explore how the surrounding condition influences the observation result. We further attempt to estimate vehicle and pedestrian flow using the sensor data. First, we attempt to identify factors that have impacts on the detection result of the sensors. Based on the field experiment in the laboratory and Gifu University campus, it is concluded that both the set height and sensor itself can influence the result. Furthermore, we got a rough detection function of vehicle flow and pedestrian flow.

***Key Words :*** *Wi-Fi packet sensor, detection efficiency, traffic flow, pedestrian flow*

## 1. INTRODUCTION

Pedestrian and vehicle volumes are one of the key criteria to evaluate the use of infrastructure and road network. If the density and flow of pedestrians and vehicles can be observed, we can give better service or save the cost, reduce air pollutions, and so on. Such kind of data can also be used for transportation planning, crowd safety, urban design, understanding human behaviour, tourist behaviour analysis and so on. Automatic counting techniques are the most promising strategy for obtaining the volume data. In the era of data-driven research, there is an interesting trend toward developing crowd-behaviour models using real-world data, and the usage of Wi-Fi packet sensor attracted the researchers' attention recently. However, many of the research just use the sensor data directly without considering the difference of detection ability of the Wi-Fi packet sensors. In fact, the detection ability of sensors may change with the surrounding conditions. Based on the above background, this research aims to explore how the surrounding condition of the Wi-Fi packet sensor influences the observation result. We further attempt to estimate vehicle and pedestrian flow using the Wi-Fi packet sensor data.

## 2. RELATED RESEARCHES

There already exist many methods to measure the amount of pedestrians such as CCTV (closed-circuit television), Bluetooth, Active badges, GPS-based systems and so on[1)]. The accuracy and reliability of two pedestrian monitoring systems under different conditions are discussed[2)]. One is utilizing downward-facing infrared depth sensor, and the other is based on a kind of visible light (RGB) camera As a conclusion, although video surveillance has a good capture rate, the method is prone to lighting conditions, viewing angles and weather conditions. Also, the cost of a video-based survey is relatively high. On the other hand analysis of mobile network GSM (Global System for Mobile Communications) log files causes strong privacy objections[3)]. The use of MAC data for tracking people has been focused recently for applying in mass events, shopping centres, airports, train stations and so on. The difference between Bluetooth and Wi-Fi wireless protocols is examined, and Wi-Fi has a more extensive operating range up to 100 meters[4)]. Similar studies can also be found[5),6),7),8)]. Judging from these researches, Wi-Fi based method does have clear advantages. The carrier of the Wi-Fi devices does not need to install any apps for collecting data, and the data outputted by the system is easy to process. Thanks for the fact that smartphones equipped with Wi-Fi modules are ubiquitous, the cost of deploying a Wi-Fi based crowd tracking system is small, and it can get continuous count. The other advantage is that it is privacy-friendly since it does not record the people's face. The pedestrian flows, wait times and counts were examined with Wi-Fi and Bluetooth sensors. The developed methods are applied to data collected at a public transportation terminal with six sensors for two months[9)]. Also, there already exists many researches focused on Wi-Fi based crowd tracking system. One study among them may be the first to use MAC address data as human movement tracking technology. They described a system using Wi-Fi detections for passively tracking smartphone clients and presented a trajectory estimation method[10)]. What's more, a system analysing pedestrian flow using Wi-Fi packet sensors was developed[11)]. Based on their research, the Wi-Fi probe request frame transmission interval is from 30 secs to 120 secs (depends on device). Then, to what extent the actual number of pedestrians can be estimated based on the Wi-Fi detection data was studied[12)]. They conducted a two-month field experiment in a shopping mall and calculated a coefficient for estimating the actual number of people within the mall by comparing the data from the Wi-Fi packet sensors with data from motion detectors. A tracking system for pedestrian flow estimation was presented and its feasibility and accuracy were investigated in a realistic scenario in a German airport[13)]. The rhythm of the campus was examined by using 20 Wi-Fi monitors to collect data for one week in the campus of the Delft University of Technology. The study focused on the user's occupation, duration of stay and moving pattern at and between the different facilities[14)]. Similarly, a one-year dataset obtained from 9 Wi-Fi tracking sensors deployed in a university campus was studied[15)].

Actually the Wi-Fi probe request is a kind of electromagnetic wave, some factors like environment's type; obstacles; distance and antenna's gain will affect the detection rate of Wi-Fi packet sensor. Other effects on wireless transmission include attenuation distortion, free space loss, noise, atmospheric absorption, multipath and refraction[16)]. Some works mentioned above tested the influence of the gain of the antenna on the Wi-Fi packet sensor observation performance, and some did not consider the basic characteristic of the sensors. They did the experiments with the assumption that the sensors have the same detection capacity. However, the observation performance of the sensors may be different even they are installed with the same antenna type considering the Wi-Fi probe request transmits characteristic.

## 3.RESEARCH QUESTION

With the development of technology, Wi-Fi technology is used almost everywhere in developed and developing countries. Also, the smart devices such as smartphones, tablets and laptop computers are very popular and almost all of them have the Wi-Fi function as a standard configuration. In this paper, as Fig.1 shows, we carried out experiments with 5 sensors (1,2,3 outdoor type; 4,5 indoor type) purchased from Japan Research Institute for Social System to test the characteristics of the Wi-Fi packet sensor and developed methods to show the potential of Wi-Fi packet data for estimating the amount of vehicles and pedestrians. The outdoor type is waterproof which has a bigger container (160 x 160 x 90 mm, 840g), whereas the indoor type has a rather small container (160 x 130 x 60 mm, 490g) and the electricity can be provided via USB socket. We first check the factors that have an impact on the detection result. Also, the ability for detecting the vehicle and pedestrian is evaluated by implementing experiments in the Gifu University campus.


**Fig.1** The photos of the Wi-Fi packet sensors

### (1) The overview of Wi-Fi packet sensors

The electronic devices, such as smartphones, tablets and computers with Wi-Fi enabled periodically transmit so called 'probe request', even when the device is not associated with a network. The probe request includes a media access control

(MAC) address which is unique per device, and thus the sensor can be used to identify the movement of its holders. It is therefore possible to count the pedestrians and vehicles through detecting the devices they are carrying. The Wi-Fi packet sensors worked as a technical device to estimate the density of the pedestrians has been developed. Fig.2 shows the overview of Wi-Fi based tracking system. An AMP sensor (Anonymised MAC Address Probe Sensor) will detect the probe request of the smart devices and upload to the cloud server after anonymising them. To protect privacy, MAC addresses are anonymised to A-MAC addresses in the sensor.

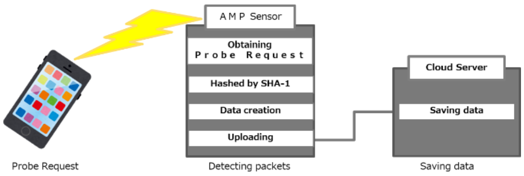


**Fig.2** Overview of Wi-Fi based tracking system

The Wi-Fi packet sensor records the following information:

**Timestamp:** Date and time when the packet is captured (accurate to second)

**Sensor ID:** Which sensor captured the packet

**AMAC:** The anonymised MAC address of the device

**OUI:** Organisationally unique identifier

**RSSI (dBm):** (Received signal strength indicator)

### (2) In-laboratory experiment

First, the devices are set on similar locations and checked the detection tendency. Then, they are put at different height locations like Fig.3 shows. The layout of the sensors was changed about every 6 days. Table 1 shows the time and location of 5 sensors and the number in the table indicate the sensor's ID (s1 to s5). We use h1, h2 and h3 to indicate Low, Middle and High location, respectively.

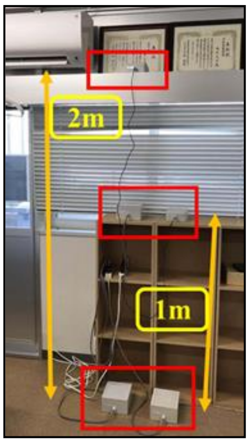


**Fig. 3** Layout of Wi-Fi sensors for height test

**Table 1** Layout of the Wi-Fi sensors

| Location / Time period | High (h3) | Middle (h2) | Low (h1) |
|---|---|---|---|
| 6/29 17:00~7/5 15:00 | s5, s4 | s2 | s3, s1 |
| 7/5 15:00~7/11 12:00 | s3 | s4, s5 | s2, s1 |
| 7/11 12:00~7/18 11:00 | s2, s1 | s3 | s4, s5 |
| 7/18 11:00~7/24 13:00 | s4 | s1, s5 | s2, s3 |
| 7/24 15:00~7/31 23:00 | s5, s3 | s4, s1 | s2 |

ANOVA (analysis of variance) has been carried out to check the statistical difference of devices themselves and height. The number of detected AMACs is aggregated by 1 hour. Before this procedure, we deleted the randomised AMAC address. There happened a power off during the height test (7/24 13:50~7/24 14:20). Therefore, we delete the data of two hours' period when the power off start and end (7/24 13:00~7/24 15:00). Table 2 shows the result. The P-value for Sensor ID and Height is smaller than 0.001, which indicates that the differences among sensors and height are statistically significant. Thus, we conclude both the sensor itself and height have impact on the detection capacity of the Wi-Fi packet sensor. To confirm this, Fig.4 shows the relationship between the average count of the sensors at a location.

**Table 2** Result of ANOVA

| Factors | DF | F value | P value |
|---|---|---|---|
| Sensor ID | 4 | 70.795*** | $<2x\ 10^{-16}$ |
| Height | 2 | 178.232*** | $<2x\ 10^{-16}$ |
| Sensor ID: Height | 8 | 3.908 ** | 0.000132 |
| ***: 0.1% significance; **:1%significance | | | |

From Fig.4 we can see when the sensors were set at the middle height the number of observations is the largest. This may be because some of the AMACs are absorbed by the ceiling and the floor. Also, the outdoor type sensor (s1, s2 and s3) have better performance than the indoor type sensors (s4 and s5). For s1 and s3, when they were set at h1 and h2, the s3 can detect more AMACs than s1; but when they were set at h3, the s1 detected more AMACs. This means the set height may have influence on the detection capacity of the Wi-Fi packet sensors. This can also be seen from Table 2. The interrelationship between sensor ID and height is also statistically significant. As Fig.5 shows, the total AMAC counts of s4 and s5 are obviously smaller than those of s1 to s3. We can conclude there really exits difference between the detection ability of the indoor and outdoor type.

Based on the in-laboratory experiment we can get conclusions that the set height and sensor itself (maybe the container of the sensor) will influence the detection capacity.

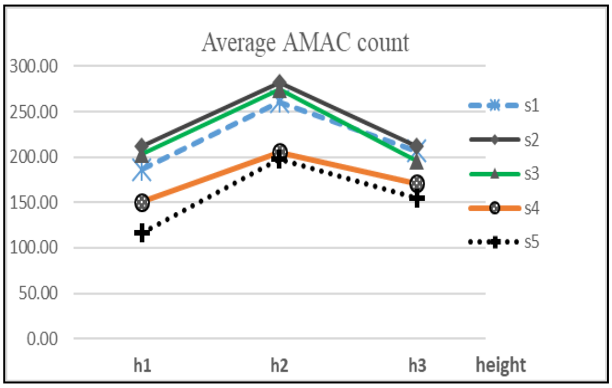


**Fig. 4** The relationship between the average count of the sensors at a location

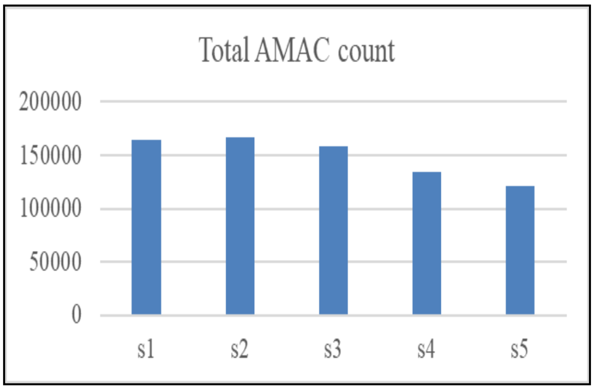


**Fig. 5** Total AMAC counts by sensor

## 4. EXPERIMENT IN UNIVERSITY CAMPUS

The second objective of this paper is to check the detection capability of sensors to quantify the number of vehicles and pedestrians. Observation experiment has been carried out in Gifu University Campus for observation of vehicles (2018/11/05 ~ 2018/12/10) and pedestrians (2019/1/11 ~ 2019/2/12). Fig. 6 represents the Wi-Fi packet sensor locations. For the observation of pedestrians, since we place the sensors where vehicles cannot enter, these areas only contain pedestrians and cyclists. At the same time, we carried out video survey to record the vehicles passing by the places and manual counting survey for recording the number of pedestrians who passed by the sensor. There are monitoring cameras at the entry gates of the university (V2, V3 and V5). We get these videos from the University and we also took video by ourselves at V1 and V4 to record the passing vehicles (2018/12/05 8:00:00 am~6:00:00 pm). For the observation of pedestrians, we set an imaginary cordon line for each sensor and count the pedestrians manually if he or she passes through the cordon line (on 2019/2/6 for P1 and P4, and on 2019/2/7 for P2, P3 and P3). We use the number of vehicles from video survey and the number of pedestrians from manually counting survey as the ground truth data. Through comparing the Wi-Fi sensor data and the ground truth data we can roughly know the detection rate of each sensor.

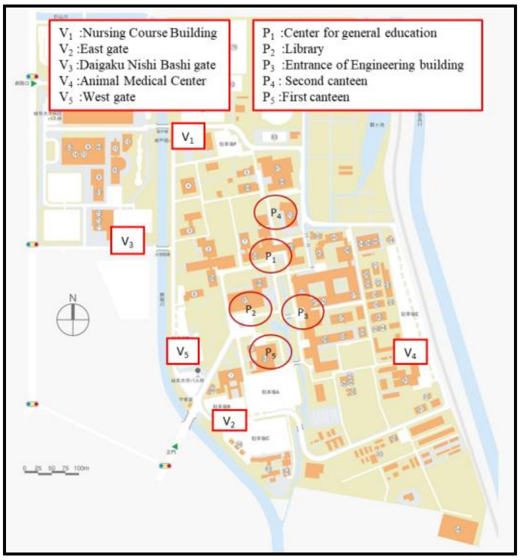


**Fig. 6** Layout of Wi-Fi packet sensors

### (1) Data pre-processing

Data pre-processing is needed since there contains noise such as the randomised AMACs and AMACs from the stationary devices like printers. The raw data provided by the sensors had been pre-processed before the analysis as follows:

**Filter 1:** Delete the randomised AMAC.

Because we can't track a device with randomised AMACs, it should be removed. The data from the Wi-Fi packet sensor contains the OUI of each AMAC, and we can identify whether the observed MAC address is randomised or not based on it. As a result, Table 3 shows the raw data of the observation of vehicles and Table 4 shows the raw data of the observation of pedestrians. The rate of randomised AMAC addresses is around 0.78 to 0.95, meaning that the valuable data are only 5 to 22%. Without considering randomised address, the result may be under/overestimated by the rate of randomised addresses.

**Filter 2:** Delete the stationary devices.

We should delete the non-mobile device since we are just interested in the movement. There may have some Wi-Fi probe requests from non-mobile devices like printers or laptops in the office. We define the non-mobile devices such that those AMACs were observed all 24 hours in a day and with the first observation time between 00:00:00 AM to 00:05:00 AM and the last observation time between 11:55:00 PM to 11:59:59 PM. V1-V5 and P1-P5 represent the sensors observing the vehicles and pedestrians. After this procedure, we can get the real AMAC data to be analysed.

**Table 3** Raw data of observation of vehicles

| Sensor | All records | Random AMAC | Real AMAC | Random rate | Time period |
|---|---|---|---|---|---|
| V1 | 35457 | 33487 | 1970 | 0.94 | 12/5 8:00~18:00 |
| V2 | 21672 | 19497 | 2175 | 0.90 | 11/26 8:00~18:00 |
| V3 | 5536 | 4761 | 775 | 0.86 | 11/26 8:00 ~18:00 |
| V4 | 7458 | 6814 | 644 | 0.91 | 12/5 8:00 ~18:00 |
| V5 | 25547 | 23521 | 2026 | 0.92 | 11/26 8:00~18:00 |

**Table 4** Raw data of observation of pedestrians

| Sensor | All records | Random AMAC | Real AMAC | Random rate | Time period |
|---|---|---|---|---|---|
| P1 | 5182 | 4178 | 1004 | 0.81 | 2/6 8:00~18:00 |
| P2 | 42924 | 39977 | 2947 | 0.93 | 2/7 8:00~18:00 |
| P3 | 25303 | 19664 | 5639 | 0.78 | 2/7 8:00~18:00 |
| P4 | 18960 | 17379 | 1581 | 0.92 | 2/6 8:00~18:00 |
| P5 | 92281 | 87801 | 4480 | 0.95 | 2/7 8:00~18:00 |

### (2) Analysis for the detection of vehicles

We further aggregated the pre-processed data by 15 minutes, and also count the vehicle from the video.

For the sensor V1 which was set at the refresh corner in Nursing Course Building, it is unavoidable to detect many AMAC records from the students nearby since many students will pass that area. However, we can identify and remove these AMACs through RSSI. Because the students are so near to the sensor than the vehicles, the signal of the AMACs from students is strong, while the signal belongs to the vehicle is weak. So, firstly we define an RSSI value which represents the vehicle, then delete the AMACs with a stronger RSSI than it. We match the data from V1 and V5 and calculate the speed of the matched AMACs because the distance between V1 and V5 is known and the time difference we can get from Wi-Fi sensor data. Generally the speed of a cyclist is 20km/h[6], therefore we select the AMACs as vehicles if its speed is greater than 20km/h. The average RSSI and median of these AMACs is -88 dBm. We regard the AMACs as the ones within the building if its RSSI is greater than -88dBm and delete them from the V1 data set. Similarly, for V4 (set in the Animal Medical Center), we should also delete the AMACs that may come from inside the building. The road condition from V2 to V4 has turns and humps, so we manually observed the time needed to cycle between V2 and V4. Based on the observation, the average time is 151s. Thus we define an AMAC is belonging to a vehicle if its time difference is less than 151s. The average RSSI and median of the matched AMACs (V2 and V4) is -86 dBm. We removed the AMACs from the V4 data set if its RSSI is stronger than -86dBm. For sensor V4 , due to the camera problem we lost 3 video data from 2018/12/5 12:30:00 pm to 2018/12/5 13:15:00, thus we have 37 data for V4 and 40 data for other 4 sensors. Through comparing the Wi-Fi sensor data and the ground truth data we can get a rough estimate function for each Wi-Fi packet sensor.

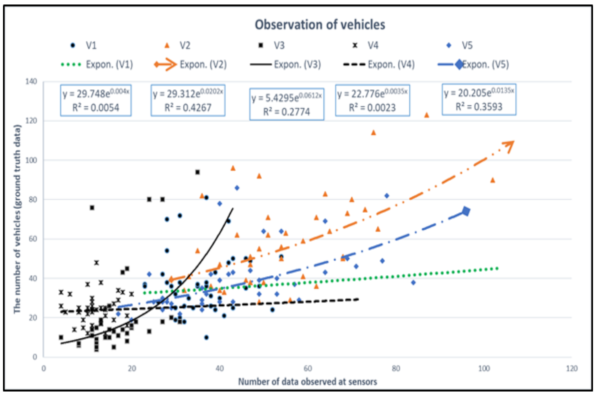


**Fig. 7** Estimating vehicle flow from Wi-Fi packet sensor data

We picked up the better fitting function based on the data set of observation of vehicles. From Fig.7 we can see, the relationship between the observed data of vehicles from Wi-Fi sensor and the ground truth data fits well with the exponential function for all 5 sensors. What's more, V2, V3 and V5 which was set at the gate of the university has relatively better fit. For V1 and V4 which was set inside the buildings far from the road, the R square values were very small. The reason may be the vehicles will stop at the gate for getting entrance permission, thus is easy to be detected by the Wi-Fi packet sensor. As for the reason for the large error for V1 and V4, we believe that this is due to the obstruction of the building. When comparing within V2, V3, and V5, the coefficient of determination for V3 is much smaller than those of V2 and V5. Considering the set location of these 3 sensors and the height test in laboratory, the V2 and V5 was set on the table in the guard room, while V3 is set on the ground because there is no guard room at Nishi Bashi gate. We can say that the location of V3 is too low and the signal may be blocked by the vehicle body, because the lower part of the vehicle is wrapped by metal.

### (3) **Analysis for the detection of pedestrians**

We counted the number of pedestrians manually as the ground truth data. We set an imaginary cordon line near the sensor, and if a pedestrian passes the line, he or she will be counted in. We selected several time periods when more pedestrians appear. The time periods varies from 15minutes to 1 hour. Through comparing the Wi-Fi sensor data and the ground truth data we can get a rough estimate function for each Wi-Fi packet sensor.

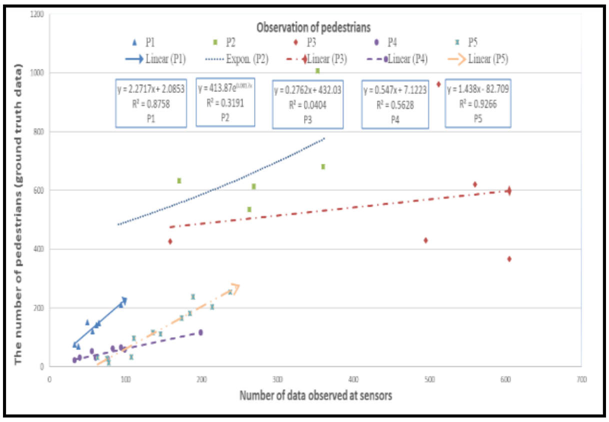


**Fig.8** Estimating pedestrian flow from Wi-Fi packet sensor data

We picked up the better fitting function based on the data set of observation of pedestrians. From Fig.8 we can see, the relationship between the observed data of pedestrians from Wi-Fi sensor and the ground truth data fits well with the linear function except P2. What's more, the coefficient of determination for P3 is very small. One of the reasons may be that P3 was set on the roof of the building, and may lose many observations due to the obstruction. For P2 and P3, the insufficient data may be another reason why they have different performance from others. From comparing the coefficient of determination for observing pedestrians and vehicles, we can say it is more accurate to count pedestrians using Wi-Fi packet sensor than vehicles except for P3.

## 5. CONCLUSION

This study attempted to find the factors influencing the detection capability of the Wi-Fi packet sensor. As a conclusion, we found that the detection capacity is not same by the sensors and also the detection rate varies by the sensor's set height. The outdoor type sensor shows better detection rate than the indoor type. What's more, in order to get more data, it is better not to place the sensor too close to the ceiling and the ground. We then studied whether we can estimate the vehicle and pedestrian flow through Wi-Fi packet sensor data. As a result, we found when detecting vehicles, the relationship between the ground truth data and Wi-Fi packet sensor data may fit well with the exponential function for all 5 sensors. While when detecting pedestrians, in some cases, the relationship between the ground truth data and Wi-Fi packet sensor data may fit well with linear function and in other cases may fit well with the exponential function. We use an average ratio between the sensor data and the ground truth data as the accuracy. The accuracy for V1 to V5 is 1.14, 1.01, 1.24, 0.71, 1.18; and for P1 to P5 is 0.44, 0.42, 0.92, 1.62, 1.86. Based on this result, we can see for the observation of vehicles it is easy overestimated, except V4 where there are very little pedestrians and cyclists passing by. For the observation of pedestrians, for P3, P4 and P5, many students will stop at these areas thus the possibility of a device to be detected will increase. While for P1 and P2, most students will pass that area instead of stop there for a long time. Therefore, we say in this case (P1, P2), it is more close to the reality for detecting the pedestrian flow. As a limitation, for the observation of vehicles, it is inescapable to include some signals from pedestrians because we can see from the video some pedestrians and cyclists also passed the sensor's detection area. Another is that not everyone carries a detectable smart device and some may carry more than one smart devices. This also has impact on the detection result. More study is needed to improve the accuracy of Wi-Fi sensor based pedestrian estimation method.

**ACKNOWLEDGMENTS:** This research is carried out as a joint research with Teikoku International Corporation.

**REFERENCES**

1) Petre, A. C., Chilipirea, C., Baratchi, M., Dobre, C., & van Steen, M. : WiFi tracking of pedestrian behavior. In *Smart Sensors Networks* pp. 309-337. Academic Press. 2017
2) Martani, C. et al. : 'Pedestrian monitoring techniques for crowd-flow prediction', Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction, 170(2), pp. 17–27, 2017

3) Giannotti, F., & Pedreschi, D. (Eds.). Mobility, data mining and privacy: Geographic knowledge discovery. Springer Science & Business Media. 2008
4) Ferro, E. and Potorti, F. : 'Bluetooth and wi-fi wireless protocols: a survey and a comparison', IEEE Wireless Communications, 12(1), pp. 12–26, 2005
5) Lesani, A. and Moreno, L. F. : 'Development and Testing of a Real-Time WiFi-Bluetooth System for Pedestrian Network Monitoring and Data Extrapolation', Transportation Research Board, 95th Annual Meeting, (January), pp. 1–13, 2016
6) Böhm, et al., : Evaluating the Usage of Wi-Fi and Bluetooth Based Sensors for Pedestrian Counting in Urban Areas. In European Transport Conference 2016Association for European Transport (AET). 2016
7) Abedi, N., Bhaskar, A. and Chung, E. :'Bluetooth and Wi-Fi MAC Address Based Crowd Data Collection and Monitoring : Benefits , Challenges and Enhancement', Australasian Transport Research Forum 2013 Proceedings 2. 2013
8) Abedi, N., Bhaskar, A., Chung, E., & Miska, M. : Assessment of antenna characteristic effects on pedestrian and cyclists travel-time estimation based on Bluetooth and WiFi MAC addresses. *Transportation Research Part C: Emerging Technologies*, *60*, pp.124-141, 2015
9) Kurkcu, A. and Ozbay, K. : 'Estimating Pedestrian Densities, Wait Times, and Flows with Wi-Fi and Bluetooth Sensors', Transportation Research Record: Journal of the Transportation Research Board, 2644(1), pp. 72–82, 2017
10) Musa, A.B.M. and Eriksson, J. : November. Tracking unmodified smartphones using wi-fi monitors. In Proceedings of the 10th ACM conference on embedded network sensor systems (pp. 281-294). ACM. 2012
11) Fukuzaki, Y. et al. : 'A pedestrian flow analysis system using Wi-Fi packet sensors to a real environment', Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct Publication - UbiComp '14 Adjunct, pp. 721–730. 2014
12) Fukuzaki, Y. et al. : September. Statistical analysis of actual number of pedestrians for Wi-Fi packet-based pedestrian flow sensing. In Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (pp. 1519-1526). ACM. 2015
13) Schauer, L., Werner, M. and Marcus, P. : 'Estimating Crowd Densities and Pedestrian Flows Using Wi-Fi and Bluetooth', Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services. 2014
14) Kalogianni, E. et al. 'Passive WiFi monitoring of the rhythm of the campus', Proceedings of The 18th AGILE International Conference on Geographic Information Science; Geographics Information Science as an Enabler of Smarter Cities and Communities, Lisboa (Portugal), June 9-14, 2015
15) Andión, J. et al. : 'Smart Behavioral Analytics over a Low-Cost IoT Wi-Fi Tracking Real Deployment', Wireless Communications and Mobile Computing, pp. 1–24, 2018
16) Beard, C., Stallings, W. and Tahiliani, M.P. : Wireless communication networks and systems, Global Edition. Pearson Education Limited. 2016


**(Received on 4th, Oct, 2019)**