A Pedestrian Flow Analysis System using Wi-Fi Packet Sensors to a Real Environment

Yuki Fukuzaki
Ritsumeikan University
1-1-1 Noji-Higashi
Kusatsu, Shiga 5258577
JAPAN
yuuki@ubi.cs.ritsumei.ac.jp

Masahiro Mochizuki
Ritsumeikan University
1-1-1 Noji-Higashi
Kusatsu, Shiga 5258577
JAPAN
moma@ubi.cs.ritsumei.ac.jp

Kazuya Murao
Ritsumeikan University
1-1-1 Noji-Higashi
Kusatsu, Shiga 5258577
JAPAN
murao@cs.ritsumei.ac.jp

Nobuhiko Nishio
Ritsumeikan University
1-1-1 Noji-Higashi
Kusatsu, Shiga 5258577
JAPAN
nishio@cs.ritsumei.ac.jp

http://dx.doi.org/10.1145/2638728.2641312

Abstract
The authors have been developing the system, which analyzes pedestrian flow using Wi-Fi packet sensors. The sensors collect Wi-Fi packet called probe request packet, which is transmitted from a smartphone to search Wi-Fi access points. In addition, the cloud storage server is running to manage observed packets centrally and to compute pedestrian flow in real time. Additionally, user movement history is vitally important and we have to pay close attention to handling that kind of data. Therefore, the system runs with an anonymization method and a cryptographic function. Some kinds of demonstration experiments were held in real environment. As a result, it was confirmed that we can analyze the rough tendency of pedestrian flow using the present system and simple analysis methods.

Author Keywords
Pedestrian Flow Analysis, Ubiquitous Computing, Sensing Network, Human Activity Recognition

ACM Classification Keywords
K.6.4 [MANAGEMENT OF COMPUTING AND INFORMATION SYSTEMS]: System Management-Quality assurance

Introduction

Since The Great East Japan Earthquake has occurred, techniques of analyzing human flow characteristics in real time have been gathering attention. Especially, it is said that analyzing human flow in urban areas is effective. Nowadays, most people carry their own smartphone. Therefore, we focus on collecting and utilizing probe request packet because the packet is transmitted periodically and has a MAC address of the smartphone. First, we developed Anonymous MAC Address Probe Sensor **(AMP sensor)**, which enables to collect MAC addresses from transmitted packets. The sensor also has upload function, and collected MAC addresses are sent to analytical server. Installing multiple sensors, we aim to analyze human flow in real time and develop a system which uses analyzed data effectively for making disaster prevention plans and taking measures to deal with natural calamities.

Most smartphones, which have Wi-Fi functions, transmit periodically probe request frame to associate with Wi-Fi access point and the frame transmission interval is from 30 sec to 120 sec (depends on device). That frame includes MAC address. Therefore, we use that as a key, because MAC address enables us to identify each device. Thus, the system can analyze the flow and distribution of persons carrying information devices. It goes without saying that movement history is a kind of personal information. Accordingly, proposed system generates secure hash value **(AMAC address)** from MAC address of an information device to assure anonymity for Wi-Fi device users.

As for the usage of AMP sensors, the number of Wi-Fi devices around an AMP sensor helps to estimating the relative number of person there. In addition, using AMAC addresses uploaded by AMP sensors installed in multiple places, pedestrians' point-to-point transit pattern can be computed. Similarly, migration velocity is also calculated. The processing of AMAC addresses enables a system to comprehend not only walking locus of pedestrians but also the trajectory data of person who travels in a car, a train and so on.

Needless to say that it is too important to analyze a pedestrian flow and a pedestrian distribution in time of disaster. However, the flow and distribution of pedestrian are also quite serviceable for urban planning and a government scheme at ordinary times. Moreover, that kinds of data become valuable information for a private enterprise such as the distribution industry and the advertising industry.

In the meantime, there is an additional approach which contrary to assure anonymity. Trajectory tracking is invaluable for watching over a child and a physically handicapped person. They are encouraged to resister personal attribute with AMAC address in advance. If so that, the system can track their real-time trajectory data as necessary. In case of an emergency, proposed system is able to function as a watching system.

Related Works

There are various approaches to analyze pedestrian tracks. For example, built-in devices on mobile phones, such as GPS receivers and accelerometers are used for recognising persons's tracks[1][2]. These kinds of approaches do not need infrastructures to collect necessary data for analyzing pedestrian trajectory, but specified applications need to be installed on persons's mobile phone. Systems developers need to consider how to spread that applications such as giving incentives for persons. Therefore, it can be said

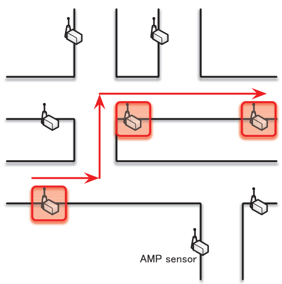


Figure 1: Pedestrian Flow Analysis as Particles

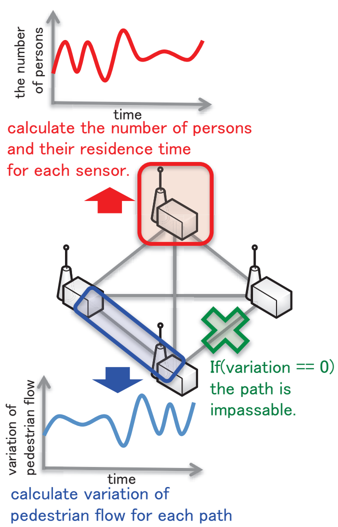


Figure 2: Pedestrian Flow Analysis as Fluid

that they are unpractical ways to collect a great number of persons' trajectory data.

In other words, if systems developers prepare an infrastructure to track pedestrians, it is unnecessary to let persons install specific applications on their mobile phones. With regard to analyzing crowd of persons, it is effective to prepare sensing infrastructure. Musa et al. [3] reported pedestrian tracking method using Wi-Fi monitor to collect probe request packets from smartphone. Viterbi algorithm is applied for estimating pedestrian trajectory. It is not required any extra configurations for smartphone. In order to improve the precision of estimating the flow and distribution of pedestrians, they take some measures. (1) emulating as popular free Wi-Fi hotspots. (2) emulating as Wi-Fi access points which have destination SSID of Directed Probe Request sent by smartphone. (3) transmitting RTS packets periodically to detected smartphone in order to receive CTS packet. Thus, they increase the number of receive packet from smartphones for improving the precision of analyzing pedestrian flow. Packet capturing concept such that is also realized using Bluetooth[4].

Moreover, there are already commercialized devices for capturing probe request packet. For example, Meshlium Xtreme[5] is a multi-protocol router made by Libelium. That provides a function to detects devices which have Wi-Fi interface or Bluetooth interface. There is also Mobility Services Engine (MSE)[6] which is provided by Cisco Systems, Inc.. MSE collects probe request packets and RSSI data from Wi-Fi access points made by Cisco Systems, Inc. through a Wireless LAN Controller (WLC). Then gathered data is used for estimating location information of mobile devices.

There are other kinds of pedestrian flow analysis methods using stereo cameras[7][8] and laser range scanners[9],[10]. With regard to using stereo camera, it is, however, quite difficult to recognize person trip data of each pedestrian by image processing because visibility of cameras are limited. For recognising individual person, it is necessary to install cameras all over the testing area or contrive a technique to identify each person on each camera. On the other hand, laser range scanner is expensive and it is also difficult to identify each person. Therefore, tracking pedestrian in large area such as a shopping mall, it costs a lot to prepare an environment for sensing person trip data.

System Concept and Architecture

Considering the introduced related works, we chose the way to prepare a infrastructure to collect Wi-Fi packet called probe request packet from Wi-Fi devices such as smartphones (It is under consideration to collect bluetooth packet in parallel), because it enables us to collect a great number of persons' data. In addition, the reason why utilizing Wi-Fi packet is that we can use MAC addresses to identify the each person, so it is easier to recognize individual person than image processing approach.

Additionally, Our proposed system has two aspects which are pedestrian flow analysis as particle (Figure 1) and pedestrian flow analysis as fluid (Figure 2). As for the former aspect, we extract a person trip data focusing on each pedestrian's trajectory. To comprehend pedestrian flow, an Origin-Destination table is generated based on person trip data. Person trip data are also used to watch over specific persons such as children and physically handicapped persons. Regarding the latter aspect, the system generates a complete graph by connecting with each sensor. The complete graph is used to analyze data such as variation of pedestrian flow, the number of resident persons and their resident time.

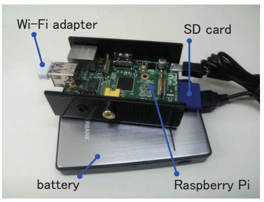


Figure 3: A prototype of AMP sensor (Type P)

Figure 4: A prototype of AMP sensor (Type N)

It is also necessary to develop a device observing probe request packets at a low price. There are some relevant devices and techniques to collect probe request packets. They are, however, not developed for the purpose of the public benefit. The feature of proposed system is not the development of AMP sensor as one of a hardware platforms. The main purpose is the spread of a function to collect and anonymize MAC addresses. In addition, we aim to build the AMP sensor network efficiently in a short period of time using existing ICT infrastructures.

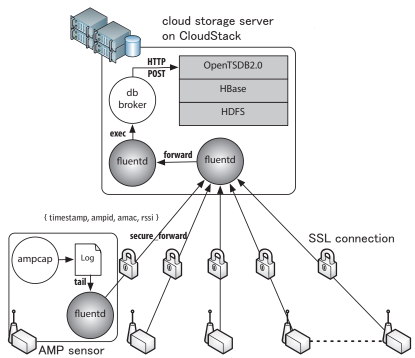


Figure 5: System Architecture

The architecture of the proposed system is shown in Figure 5. To handle a large amount of data, the cloud storage server should provide scalability and high-speed data processing capability. That is why that a server is built on a cloud server and with the processing engine of distributed database. The data captured by AMP sensors are encrypted and communication path are also secured.

AMP Sensors

Considering the installation in multiple places, It is desirable that the AMP sensor is developed as a small-sized device and developed at a low price. In addition, it is also desirable that the only thing to boot and run the AMP sensor is just supplying electricity. In case of mass production, the objective is developing the AMP sensor under one hundred dollars. An AMP sensor has an anonymization method and a cryptographic function to prevent the leak of personal information.

Hardware

Table 1: Specifications for AMP Sensors

Item	Type P	Type N
form factors	Raspberry PI Type B	Intel NUC (DC3217IYE)
size	100 (wide) × 64 (depth) × 30 (height) mm	116.6 (wide) × 112 (depth) × 39 (height) mm
CPU	ARM1176JZF-S 700 MHz	Intel Core i3 3217-U
main memory	512 MB	SO-DIMM 8 GB DDR3 (1600 MHz)
storage	8 GB	128 GB SSD (mSATA)
network	wired network: 10/100 Mbps ethernet (RJ45) wireless network: Logitec LAN-W150NU2AW	wired network: Integrated Intel PRO 10/100/1000 wireless network: Intel Centrino Advanced-N 6235
power supply	5 V, 3.5 W AC adapter	19 V, 65 W AC adapter

Regarding the hardware of the AMP sensor, Wi-Fi adapter which runs as monitor mode and PC are needed. At this moment, two different prototypes are developed; one is an AMP sensor (Type P) based on Raspberry PI Model B shown in Figure 3 and the other is AMP sensor (Type N) based on Intel NUC Platform shown in Figure 4. Specifications for each AMP Sensor are described in Table 1.

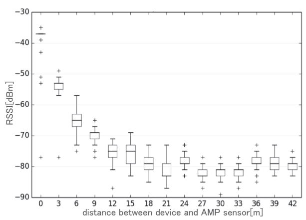


Figure 6: Inside a Building

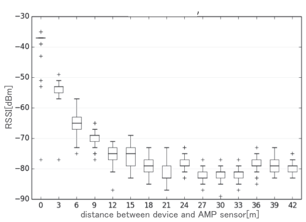


Figure 7: Outside a Building

Software
Operating system (OS) of an AMP sensor (Type P) is Linux-based OS named Raspbian3.6.1+. OS of an AMP sensor (Type N) is Ubuntu12.04LTS. At this moment, when AMP sensor observed a probe request packet, the MAC address is hashed by SHA-1(160bit) and the sensor archives hashed value, RSSI, observed time and observation point as a log. The archived logs are periodically uploaded to cloud storage server using a fluentd. An OpenSSL library is used for hashing MAC address and a libcap is used for capturing packets. AMP sensor software is written in C and C++.

Basic Characteristic
We made an experiment to measure the basic characteristics of AMP sensor. We verified RSSI attenuation by distance between a smartphone and an AMP sensor. In this experiment, we prepared Android OS smartphone named Galaxy Nexus (indoor environment : Figure 6, outdoor environment : Figure 7). An RSSI value is clearly high within around 15 m. If the distance is over 15 m, it is difficult to say that there is a proportional connection between an RSSI value and a distance. RSSI also depends on the environment and the way of holding a smartphone. Experimental results regarding packet transmission interval is shown in Table 2. As for each device, we confirmed that each device transmits probe request packets periodically. However, the transmission interval depends on the device. In addition, there are some devices which do not transmit probe request packets.

Demonstration Experiment

We made experiments in a real environment. The purpose is to verify whether a simple pedestrian flow analysis method is applicable or not in a real environment.

Experiment in ActiveLab
This experiment was made at ActiveLab in GRAND FRONT OSAKA. The Experimental period is from January/29/2012 16:00 to January/31/2012 18:00. In addition, we made reference points and an experimenter moved between designated points at a specified time while holding a smartphone which has an already - known AMAC address (from January/29/2012 18:00). The person who holds that device made a one-minute stay at each reference point. We show an experimental result regarding reference points. A floor map, installation locations of AMP sensors (J1 to J6) and reference points (R1 to R20) are shown in Figure 8. A list of reference point names and the start time of staying is shown in Table 3. During the observation in regard to underlined reference points, smartphone was put just above underlined reference points.

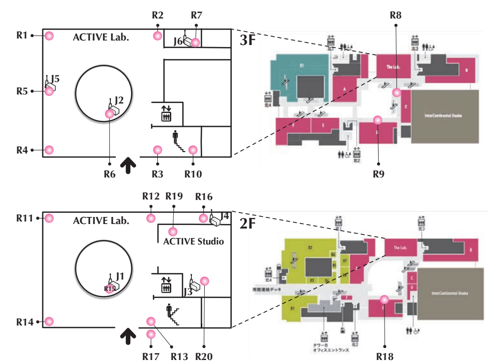


Figure 8: ActiveLab Floor Map

Table 2: The Transmission Intervals of Probe Request

OS	device name	maximum interval(s)	remarks
Android	Galaxy Nexus	247.7	repeat a regular set (30s×4times, 60s×4times, 120s×4times, 240s×4times)
	Nexus 4	14.8	approximately 15s interval
	Xperia Tablet S	—	can't observe any probe request packets
	Eee Pad TF201	—	can't observe any probe request packets
iOS	iPhone4S	481.2	approximately 480s interval
	iPad 1G	483.1	approximately 480s interval

Table 4: A Result of Sampling at Reference Points

name	answer	result
R5	J5	J5
R6	J2	**J5**
R7	J6	J6
R15	J1	J1
R16	J4	J4
R20	J3	J3

Table 3: A List of Reference Points

name	observing start time	name	observing start time
R1	18:18	R11	18:41
R2	18:20	R12	18:43
R3	18:22	R13	18:45
R4	18:24	R14	18:47
R5	18:26	R15	18:49
R6	18:28	R16	18:51
R7	18:30	R17	18:53
R8	18:33	R18	18:55
R9	18:35	R19	19:05
R10	18:39	R20	19:07

Experiment Result Regarding Reference Point

Figure 9 shows time-series plot of AMP sensor ID which observed target AMAC address. Collected data are resampled at one minute intervals and we consider that smartphone exits near the AMP sensor whose RSSI is max value within a resampling interval. Vertical lines mean observing start timings on each reference point. In addition, a solid line among them means the case smartphone was put just above a reference point. With regard to each reference point showed as solid line, the comparison of ground truth AMP sensor ID and estimated AMP sensor ID which is calculated from RSSI value is shown in Table 4. Regarding reference point R6, the answer is J2, however the estimated AMP sensor ID is J5. Although there were a little discrepancies, we could obtain expected results. Therefor RSSI can be used to estimate users location.

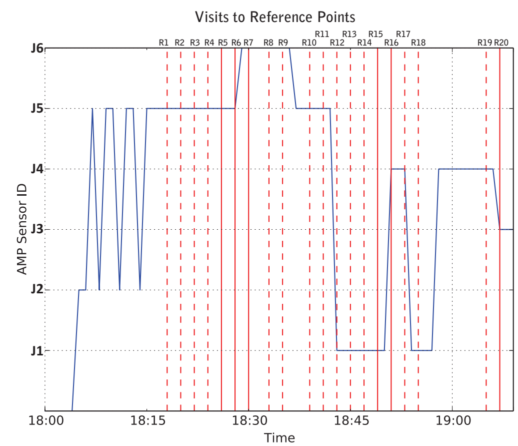


Figure 9: Visits to Reference Points

Experiment in Osaka Electro-Communication University

This experiment is made during the graduation work exhibition called Nawaten in Osaka

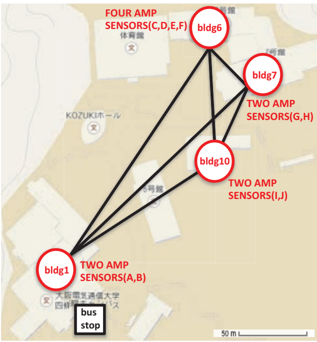


Figure 10: Osaka Electro-Communication University Campus Map

Table 5: Classification of Observed Devices

type	quantity of devices
portable	320
fixed	385
undetermined	572
Σ	1277

Table 6: Mean Value of Resident Time

building number	mean value of time(min)
building1	55.4
building6	40.0
building7	36.4
building10	39.2

Electro-Communication University. The experimental period was from February/8/2012 11:00 to February/9/2012 17:00. Figure 10 shows campus map of Osaka Electro-Communication University and installation locations of AMP sensors. There is a bus stop near the building1, and most participants used buses because of site condition. In addition, participants of this exhibition had to sign in at the reception desk in building1. That is why who comes into university and who goes away from university was expected to walk by building1. Graduation works were displayed in each building. The graduation work exhibition was held from 11:00 to 17:00 in both days. There were also special events, one was social gathering held from 15:00 to 17:00 on the first day in building6, the other was commendation ceremony from 15:00 to 16:15 on the second day in building7. Almost all participants attended above events. During this experiment, almost all the AMP sensors could not observe correct RSSI value. That is why we did not use RSSI values in this analysis. For example, if the probe request transmitted by a device is observed Sensor A or B, we consider that the device is observed in building1. Resampling interval is 2 minutes, and the smartphone is regarded that it is in the building which observed the smartphone's AMAC address with highest frequency within the resampling interval.

Classification of Wi-Fi Device

Not only portable devices such as smartphones but also fixed devices such as Wi-fi base stations transmit probe request packets. Therefore, we need to classified the type of Wi-Fi devices. We divided them into three types (**portable devices, fixed devices and undetermined devices**). The devices which were observed in at least two buildings were regarded as portable devices. The devices which were observed all through the night were regarded as fixed devices. Concretely, we devided 9 hours from February/8/2012 21:00 to February/9/2012 6:00 into 9 periods, and devices observed in every period were treated as fixed devices. The devices satisfying both portable device condition and fixed device condition were regarded as fixed devices. After that, remaining devices were treated as undetermined devices which were difficult to decide as portable devices or fixed devices. The classification result is shown in Table 5. The unique number of observed devices was 1277. The number of portable devices was detected 320 and that of fixed device was 385. 572 devices were regard as undetermined devices.

Generation of Pedestrian's Trajectory Data

As for analyzing a pedestrian flow, it is the first step to generate pedestrian's trajectory data. Figure 11 shows the process of generating trajectory data regarding a portable device.

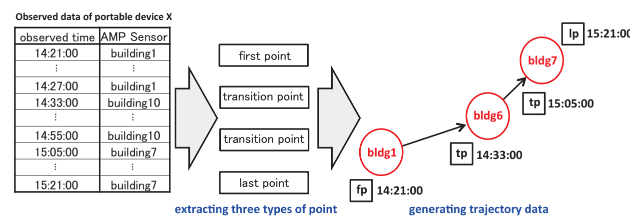
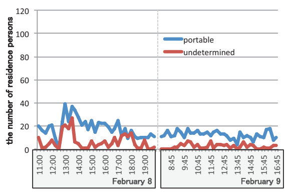
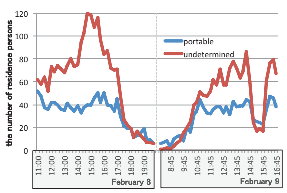


Figure 11: Generation of Trajectory Data

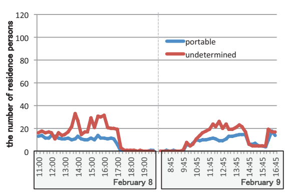
Three types of points (**first point, transition point and last point**) are extracted. The point, where the portable device X was observed at first, is regarded as the first point. The point, which changes observed point, is treated as a transition point. The other, where the portable device X was observed the last, is regarded as the last point. After extracting each point, pedestrian's



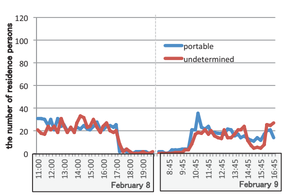
(a) building1



(b) building6



(c) building7



(d) building10

Figure 12: The Number of Portable/Undetermined

trajectory data is generated by connecting those points in chronological order.

Origin-Destination Table
Using the transition points, the system can know where the pedestrian comes from and where he/she goes. For generating a Origin-Destination table, the data with regard to the relation between origin and destination were aggregated. Table 7 is Origin-Destination table generated using extracted transition points.

Table 7: Origin-Destination Table

		Destination				Σ
		bldg1	bldg6	bldg7	bldg10	
Origin	bldg1	0	131	11	27	169
	bldg6	90	0	416	272	778
	bldg7	11	402	0	15	428
	bldg10	40	266	9	0	315
	Σ	141	799	436	314	1690

Calculation of Resident Time
At least one transition point is generated, resident time can be roughly calculated. The difference time between the transition point and the previous transition point is considered as resident time. Mean value of resident time regarding this experiment is shown in Table 6.

Counting the Persons Comes into/Goes out of University
Considering the locations of bus stop and a reception desk, almost all participants may be considered to walk by or drop at building1. Therefore, we detect a time when the person comes into university or goes out of university using trajectory data. If the first point is generated in building1, that time is regarded as a time when the person comes into university. On the other hand, If the last point is generated in building1, that time is regarded as a time when the person goes out of university. The each number of both persons at 30-minute intervals is shown in Figure 13. A blue line is the number of persons who come into university, and a red line is the number of persons who go out of university.

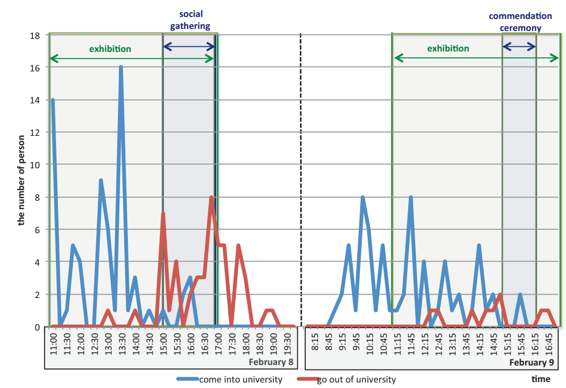


Figure 13: The Number of Persons Come into University/Go out of University

According to the results, most persons who come into university is tend to appear until 13:00 o'clock, and most persons who goes out of university is tend to leave around 17:00 o'clock.

Management Of Undetermined Devices
The way to treat undetermined devices should be considered. The comparison regarding the numbers of portable devices and undetermined devices at 30-minute intervals are shown in Figure 12. According to the results, the waveforms regarding portable devices and that of undetermined devices look quite similar. An undetermined

may be considered as a portable device. At least in this analysis, these undetermined devices were treated as portables devices.

The Number of Resident Person in Each Building
The number of resident persons at 30-minute intervals with regard to each building is shown in Figure 14. The time of the exhibition and that of the special event are also shown. In addition, the number of resident persons are estimated using information of portable devices and undetermined devices.

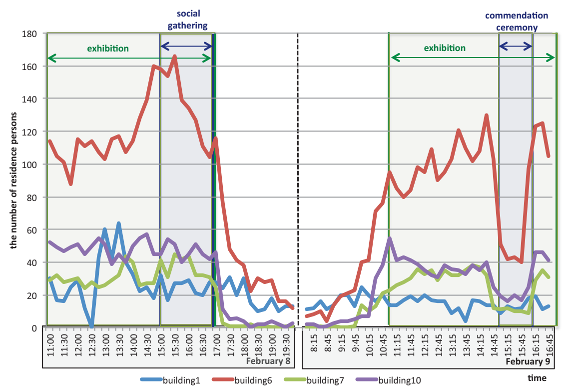


Figure 14: The Number of Resident Persons Based On Portable and Grey Devices

On the first day, a social gathering was held from 15:00 to 17:00 in building6. Therefore, the number of the resident person in building6 peaked during the event period. On the other hand, a commendation ceremony was held in building7 from 15:00 to 16:15 on the second day. almost all the participants took part in the event. The number of resident person in building7, however, decreased. Furthermore, the number of resident persons in each building decreased sharply at the event duration. We guess due to the difficulty of observing probe request packets from smartphone in the event hall because of the sensor location. Therefore, only the considerable decline in each building were confirmed from the result. In addition, the number of the resident persons was tend to increase until 13:00 o'clock, and that number was tend to decrease after ending time of the exhibition.

Conclusion

Most smartphones, which have Wi-Fi functions, usually sends probe request frame to associate with a Wi-Fi access point and the frame transmitting interval is from 30 sec to 120 sec (depends on device). A probe request frame includes MAC address, therefore, we collect each MAC address as a hashed value to analyze the pedestrian flow. We described the result of some demonstration experiments held in a real environment. We confirmed that we can analyze the rough tendency of the pedestrian flow using the present system and simple analysis methods. As for future issues, it is necessary to generate more accurate person trip data because it becomes fundamental information in our analysis. Moreover, we are planning to implement a system which can analyze a pedestrian flow in real time on a server, as soon as sensor data are uploaded.

Acknowledgements

We received generous support from Strategic Information and Communications R&D Promotion Programme (No.132307011) organised by Ministry of Internal Affairs and Communications.

References

[1] M. Wirz, T. Franke, D. Roggen, E. Mitleton-Kelly, P. Lukowicz, and G. Troster. Inferring crowd

conditions from pedestrians' location traces for real-time crowd monitoring during city-scale mass gatherings. In *Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2012 IEEE 21st International Workshop on*, pp. 367–372, June 2012.

[2] Seung Hyuck Shin, Chan Gook Park, and Sangon Choi. New map-matching algorithm using virtual track for pedestrian dead reckoning. *ETRI Journal*, Vol. 32, No. 6, pp. 891–900, 12 2010.

[3] A. B. M. Musa and Jakob Eriksson. Tracking unmodified smartphones using wi-fi monitors. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, SenSys '12, pp. 281–294, New York, NY, USA, 2012. ACM.

[4] M Abbott-Jard, H Shah, and A Bhaskar. Empirical evaluation of bluetooth and wifi scanning for road transport. *Australasian Transport Research Forum (ATRF), 36th, 2013, Brisbane, Queensland, Australia*, p. 14, 2013.

[5] Libelium. Meshlium xtreme datasheet (document version:v4.2 - 11/2003).

[6] James Little, BrendanO ' Brien. A technical review of cisco ' s wi-fi-based location analytics white paper.

[7] Janne Heikkilä and Olli Silvén. A real-time system for monitoring of cyclists and pedestrians. *Image and Vision Computing*, Vol. 22, No. 7, pp. 563–570, 2004.

[8] Q. Cai and J. K. Aggarwal. Tracking human motion using multiple cameras. In *In Proc. of the 13th International Conference on Pattern Recognition*, pp. 68–72, 1996.

[9] Huijing Zhao and R. Shibasaki. A novel system for tracking pedestrians using multiple single-row laser-range scanners. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, Vol. 35, No. 2, pp. 283–291, March 2005.

[10] K.O. Arras, S. Grzonka, M. Luber, and W. Burgard. Efficient people tracking in laser range data using a multi-hypothesis leg-tracker with adaptive occlusion probabilities. In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pp. 1710–1715, May 2008.