

Estimating People Flow and Crowdedness for Various Urban Environments Based on BLE Signal Sensing: Practical Studies

Yuki MATSUDA^{†,††,†††a)}, Hirohiko SUWA^{††}, *Members*, Kotaro HAYASHI[†], Taito YOSHIMURA[†], Arata YOSHIHARA^{††}, *Nonmembers*, and Ismail ARAI^{††}, *Senior Member*

SUMMARY To realize Society 5.0, the construction of urban digital twins is an urgent task. Among the various concern, human mobility is one of the most critical aspects that has attracted significant attention, and numerous approaches have been proposed so far. For example, the methods using CCTV cameras or LiDAR could perform high quality estimation, but they require the acquisition of sensitive data, e.g., video footage or gait data, and pose challenges of social acceptability due to the perception of surveillance. Therefore, this study aims to develop methods to estimate the number of people (crowdedness level) stay in various public spaces and the movement between places in the city (people flow) by utilizing WiFi and/or BLE (Bluetooth Low Energy) signals emitted by personally owned smartphones. In particular, this paper introduces a method for estimating crowdedness levels and people flow based on BLE advertising packet data (including RSSI and random addresses), and the five practical studies in the real-world settings.

key words: Society 5.0, Smart city, crowdedness level, people flow, Bluetooth Low Energy

1. Introduction

In anticipation of the Society 5.0 era, the realization of urban-scale digital twins that map the physical space we live in (physical space) to the digital world (cyber space) is highly expected, enabling various analyses and simulations. For instance, estimating and predicting the number of people (crowdedness level) stay in various public spaces and the movement between places in the city (people flow) could contribute to urban efficiency and commercial strategy planning. According to the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) guide “Guidelines for the Utilization of people flow Data to Solve Regional Issues” [1], the primary utilization usecases of people flow data include current status assessment, evaluation of policy effectiveness, and basic data for forecasting. By combining people flow data with other types of data, it is anticipated to lead to Evidence-Based Policy Making (EBPM), consequently, the demand for people flow data is expected to increase further in the future.

However, manually measuring crowdedness levels and people flow is labor-intensive, and acquiring comprehensive crowdedness data across urban areas and individual mobility log data is extremely challenging. To address these challenges, various ICT-based approaches have been proposed. For example, methods that estimate people flow using AI cameras or 3D sensors attempt to track movement of people by performing person re-identification based on its appearance and/or gait information obtained from sensors. However, these methods face challenges in terms of high equipment installation and operational costs, as well as social acceptance issues due to the perception of surveillance. To address this issue, other approaches have been proposed to estimate crowdedness levels and people flow using WiFi and Bluetooth packets [2], [3]. However, these methods do not adequately address the challenge posed by the recent adoption of MAC address randomization for privacy protection.

From these backgrounds, this study aims to estimate crowdedness levels and people flow in various urban spaces by leveraging advertising packets of BLE (Bluetooth Low Energy) that include randomized MAC addresses (random addresses). We have proposed various methods suitable for different environments to achieve this goal [4]–[12] (Fig. 1). In this paper, we survey the basic characteristics of advertising packets, and introduce the overview of crowdedness level and people flow estimation methods and the five practical studies conducted in the real-world settings.

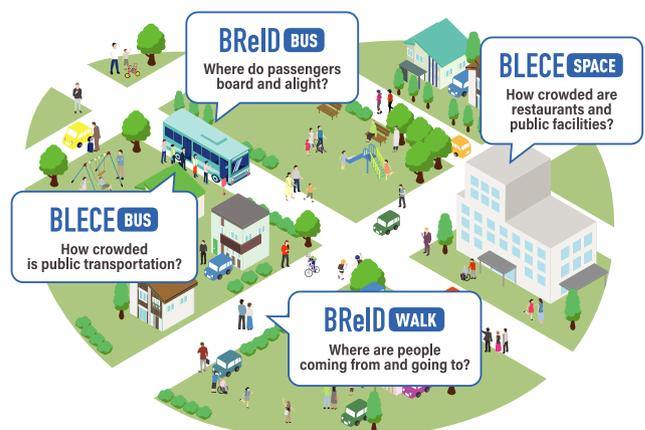


Fig. 1 The methods of estimating people flow and crowdedness level using BLE advertising packet data.

Manuscript received March 15, 2025.

Manuscript revised May 20, 2025.

Manuscript publicized July 1, 2025.

[†]Faculty of Environmental, Life, Natural Science and Technology, Okayama University, Okayama-shi, 700-0082 Japan.

^{††}Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma-shi, 630-0192 Japan.

^{†††}RIKEN Center for Advanced Intelligence Project (AIP), Tokyo, 103-0027 Japan.

a) E-mail: yukimat@okayama-u.ac.jp

DOI: 10.23919/transcom.2025CEI0006

2. Related Work

Here, we introduce related work about methods of estimating crowdedness levels and people flow.

2.1 Crowdedness Level Estimation

Crowdedness level estimation has long been the subject of academic research, and several commercial services are already available.

First, as existing services that provide crowdedness level information in urban environments, we can mention Mobile Spatial Statistics by NTT Docomo [13] and Kompreno by Agoop Inc. [14]. These services collect location information transmitted from GPS-enabled mobile devices, including smartphones, with the user's consent through applications provided by each company. They visualize population estimates on a mesh level of fixed sizes, such as 125 m or 250 m squares. However, it is challenging to estimate the crowdedness level in specific locations.

In the field of computer vision, estimating crowdedness level and people flow through image analysis using cameras has been a long-standing research focus. Sindagi et al. [15] proposed a method for estimating crowd density and number of people by fusing multiple convolutional neural network (CNN) components that incorporate both global and local contextual information of crowd images. Liu et al. [16] proposed an end-to-end trainable deep architecture called Context-Aware Network, which combines features obtained from multiple receptive field sizes to estimate number of people in crowded scenes. Some studies and services also utilize laser imaging, detection, and ranging (LiDAR) technology. SICK AG [17] provides a system for counting the number of people entering or leaving an area using 3D-LiDAR sensors.

Furthermore, several methods have been proposed to use various sensors equipped on mobile devices. Kannan et al. [18] proposed a participatory crowdedness level estimation system where sound signals emitted from a smartphone are received by another user's device, and acoustic characteristics are analyzed. Nishimura et al. [19] proposed a method to estimate the smoothness of pedestrian flows using accelerometer data and ambient sound data collected from smartphones. Moustafa et al. [20] proposed a method to estimate crowdedness level of railway station by analyzing passengers' behavior using smartphone motion sensors and capturing ambient sound characteristics via the microphone.

Methods to estimate crowdedness level in public transportation have also been proposed. Some methods use cameras installed inside transportation vehicles to estimate crowdedness levels [21], [22]. For instance, Song et al. [21] proposed a system for counting the number of passengers using images from surveillance cameras. Although these approaches enable relatively accurate estimation by directly recognizing images of people, the installation and operation of cameras involve high costs and may pose social acceptance issues due to privacy concerns.

There are also methods that utilize radio waves such as Wi-Fi and BLE. Handte et al. [23] proposed a method for estimating the number of passengers on buses by counting the MAC addresses of passengers' mobile devices connected to the bus Wi-Fi access point. Hidayat et al. [24] utilized GPS and Wi-Fi scanners to detect MAC addresses of bus passengers for estimating the number of passengers. Maekawa et al. [25] proposed a method that scans surrounding Bluetooth signals using passengers' smartphones to estimate crowdedness levels. Weppner et al. [26] proposed a method to estimate crowd density by collecting BLE signals received by mobile devices of people moving in a given environment and counting the number of nearby BLE devices. Takahashi et al. [27] combines an overhead camera and a Wi-Fi scanner that detects probe requests from people's mobile devices for estimating crowdedness of bus stations.

2.2 People Flow Estimation

Various methods have been proposed for people flow estimation.

Wang et al. [28] proposed a person re-identification method for human tracking using multiple cameras. By leveraging visual features from camera footage and spatial-temporal constraints based on camera placement and recording time, their method enables highly accurate tracking of individuals. Numerous other studies also address person re-identification methods using multi-camera footage. Nagata et al. [29] proposed a people flow estimation method using 3D passage sensors. These sensors are installed at all entrances and exits of the target area to collect information such as height, walking speed, passage time, and passage direction, which are then matched across observation points to estimate people flow. Yamaguchi et al. [30] proposed *Hitonavi*, a method that uses LiDAR to estimate the location and behavior of individuals and human crowds. However, these methods face high equipment and operational costs, making practical implementation difficult for the whole city.

With the widespread adoption of smartphones, research on people flow estimation using Wi-Fi and Bluetooth signals emitted from such devices has been conducted. Versichele et al. [31] proposed a proximity-based Bluetooth tracking method using multiple Bluetooth scanners to collect MAC addresses at various points during large-scale events. Murai et al. [32] proposed a method to estimate people flow by capturing non-randomized MAC addresses using multiple Wi-Fi scanners. When the same MAC address is detected at different scanner locations, it is considered as a moving device, and origin, transfer, and destination points are identified to generate a movement trajectory. However, these methods do not address recent privacy protection measures, such as randomized MAC addresses.

2.3 Position of This Study

Although many methods for estimating crowdedness levels and people flow in urban spaces have been proposed, chal-

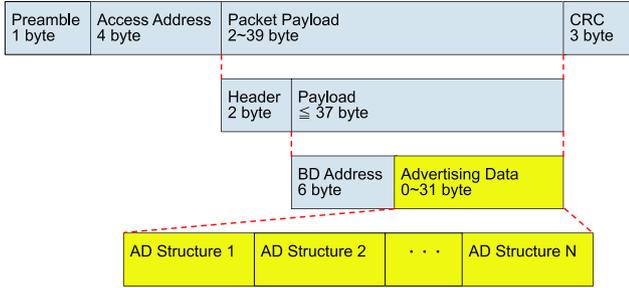


Fig. 2 BLE advertising packets format.

lenges remain regarding privacy preservation, implementation, and operation costs, making it difficult to widely apply these methods in real-world scenarios.

To address these issues, the authors have previously proposed a series of studies focusing on the utilization of BLE advertising packets with randomized MAC addresses for various applications, including crowdedness estimation in different types of spaces and people flow estimation between locations [4]–[12].

This paper provides a set of practical case studies that explore the potential of utilizing BLE advertising packets for crowdedness and people flow estimation. By systematically introducing these examples across different urban spaces and use cases, the paper highlights the applicability and versatility of BLE-based sensing approaches, contributing to the discussion on how such methods can be leveraged for human mobility analysis.

3. Preliminary Investigation of Advertising Packet Transmission

While some basic information about advertising packets is available from official resources such as the Bluetooth SIG specifications, the detailed structure of the packet contents and the actual conditions under which packets are transmitted (e.g., transmission frequency, timing, and variations across device models or operating system versions) are not fully disclosed. Furthermore, these behaviors may differ depending on the device model, OS version, or user settings. Since such variations directly affect the feasibility and accuracy of BLE-based crowdedness and people flow estimation, it is essential to experimentally investigate how these packets are actually transmitted by real devices.

For these reasons, we conducted an empirical investigation using actual devices to verify and characterize the behavior of advertising packets (Fig. 2) that contain randomized MAC addresses (random addresses). In this study, we focused on widely used iOS and Android devices.

3.1 Advertising Packets Transmitted by iOS Devices

First, we investigated the packets emitted by iOS devices [33], [34]. As a BLE advertising packet collection

device (scanner), we used a Raspberry Pi 4 Model B[†] equipped with a Bluetooth 4.0+EDR/LE Class1 compatible USB adapter (BUFFALO, BSBT4D100). We used five iOS devices for the investigation: iPhone SE 2nd Generation (iOS 16.11), iPhone XR (iOS 17.5.1), iPhone 13 (iOS 16.2), iPhone 15 Pro Max (iOS 17.0), and iPhone 13 mini (iOS 15.3.1). Additionally, we used two Apple Watch devices that may emit similar BLE packets: Apple Watch Series 5 (watchOS 10.3) and Apple Watch Series 9 (watchOS 10.3). To collect advertising packets, we used the Python library bluepy^{††}, and placed both the iOS devices and the BLE scanner inside a radio anechoic box (MICRONIX, MY1510) to capture the data.

The presence or absence of packet detection, as well as the contents of detected packets, are shown in Table 1. It was found that Apple products always emit data starting with `0x4c0010` as Manufacturer Specific Data (MSD) included in the Advertising Data (Fig. 3). The first four digits of the MSD (e.g., `0x004c`) are referred to as the Company ID, which identifies the manufacturer (`0x004c` corresponds to Apple Inc.).

The data actually acquired has the upper 2 bytes in a swapped order (for Apple devices, it is obtained as `0x4c00`). While the upper 2 bytes are specified by the BLE protocol, the subsequent data is determined at the discretion of the manufacturer, resulting in variations between devices and manufacturers. If not publicly disclosed, third parties cannot determine the contents.

Next, to investigate how frequently the BD address is randomized and how the packet contents (MSD, Tx Power Level) change over time, we placed an iPhone 15 Pro Max (iOS 17.5.1) and a BLE scanner inside the radio anechoic chamber and collected data continuously for 20 hours.

As a result of the investigation, it was found that the iPhone 15 Pro Max continuously transmits two types of packets, one always starting with `0x4c0010` and the other with `0x4c0012` as MSD. During the 20-hour period, the number of BD address changes was 39 times for packets starting with `0x4c0010` and 74 times for packets starting with `0x4c0012`.

When taking a 5-hour segment from the 20-hour period, as shown in Fig. 4, it was observed that the duration of each BD address varies (`0x4c0010` – maximum: 39 minutes 32 seconds, minimum: 19 minutes 58 seconds, average: 31 minutes 36 seconds; `0x4c0012` – maximum: 31 minutes 51 seconds, minimum: 1 minute 6 seconds, average: 16 minutes 13 seconds). Additionally, the Tx Power Level could only be extracted from packets starting with `0x4c0010`, and its value remained constant at `07` throughout the observation.

3.2 Advertising Packets Transmitted by Android Devices

Next, we investigated the packets emitted by Android devices. For collecting BLE advertising packets, we used a

[†] <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/>

^{††} <https://github.com/IanHarvey/bluepy>

Table 1 Advertising Packet Transmission Status (iOS).

Device (OS ver.)	Usage over 3 months	Packet Transmission	Manufacturer Specific Data
iPhone SE 2nd-Gen. (iOS 16.11)	✓	✓	4c0010053618a266a5
iPhone XR (iOS 17.5.1)	✓	✓	4c0010070f1f7a89710048 4c0012020000 4c001608004781ead547e95c
iPhone 13 (iOS 16.2)	✓	✓	4c0010063b1a2c6d3068 4c0012020001
iPhone 15 Pro Max (iOS 17.0)	✓	✓	4c0010062619e50f4328 4c0012190093eeb5f8bdd073e8028328a54433c8a21a59b60d3f060300
iPhone 13 mini (iOS 15.3.1)	-	-	
Apple Watch Series 5 (watchOS 10.x)	✓	✓	4c0010020100
Apple Watch Series 9 (watchOS 10.3)	✓	✓	4c0010052d187b3bf0 4c0012190015d300edeb44ecd0857dc26b78178fd91e20cb8d2000200

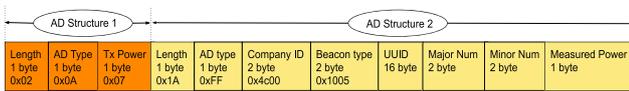


Fig. 3 Advertising data format example of iOS devices.



Fig. 5 Advertising data format of android devices.

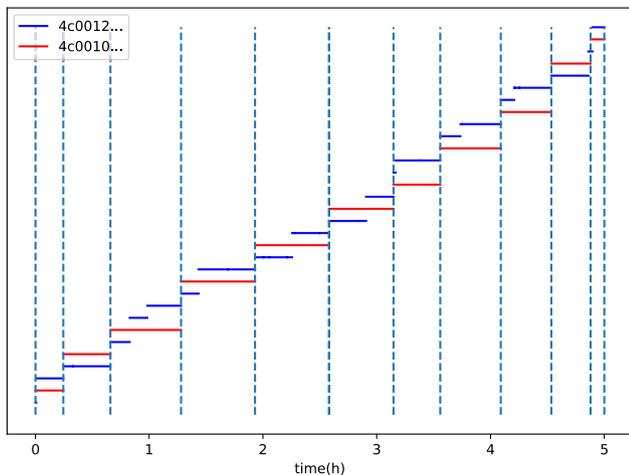


Fig. 4 BD address changes (dashed lines indicate address change points).

Table 2 Examples of observed service data.

Service Data Example	Byte Length
4a1723593352381132bd6b4d886b65447c9be4e10...	27
11012ba6cdb4	6

Unlike the Advertising Data of iPhones, Android devices include the Data Type “Service Data - 16-bit UUID” in their Advertising Data. This field consists of four items: length, type, UUID 16, and Service Data. All the Android devices used in the experiment were confirmed to emit packets following this format. According to the Bluetooth SIG Assigned Numbers document [35], the Type field is 0x16 and the UUID 16 field is 0xfe f3 respectively for packets emitted from these devices (i.e., the OS provided by Google LLC).

The Event Type of the packets observed was either Scannable Undirected Advertising or Scan Response. During active scanning, when a scannable advertising packet is observed, the scanning device sends a scan request packet (hereafter, Scan Request). The Android device that receives this packet is assumed to send a scan response packet (hereafter, Scan Response). Therefore, Scannable Undirected Advertising and Scan Response can be considered as a pair.

We mainly identified two types of pairs from Android devices. Specifically, some packets had Service Data in the Scan Response, while others had Service Data in the Scannable Undirected Advertising. Examples of each Service Data field are shown in Table 2. The Service Data field in Scan Response has a data length of 27 bytes, where bytes 1–3 and bytes 8–9 did not change during the experiment. On the other hand, the Service Data field in Scannable Undirected Advertising has a data length of 6 bytes, where bytes 1–2 remained unchanged during the experiment.

Packets with a Service Data length of 27 bytes were observed on all tested devices. During the experiment, Ser-

Raspberry Pi 4 Model B equipped with a Bluetooth 4.0 compatible USB adapter (TP-Link, UB400) as the scanning device. As Android devices, we used four models: Google Pixel 4a (Android 13), Google Pixel 6 Pro (Android 14), AQUOS sense3 (Android 11), and ROG Phone 5s (Android 13). All devices were signed in with the same Google account during initial setup, and all other settings were left at their default values. The scanning device was placed in close proximity to each Android device, and packet collection was conducted for one hour. To capture packets, we used hcitool version 5.66[†] and btmon version 5.66^{††}. Active scanning was performed during the collection process.

Among the packets obtained from Android devices, the structure of the Advertising Data part is shown in Fig. 5.

[†] <https://github.com/bluez/bluez/wiki/hcitool>

^{††} <https://github.com/bluez/bluez/wiki/btmon>

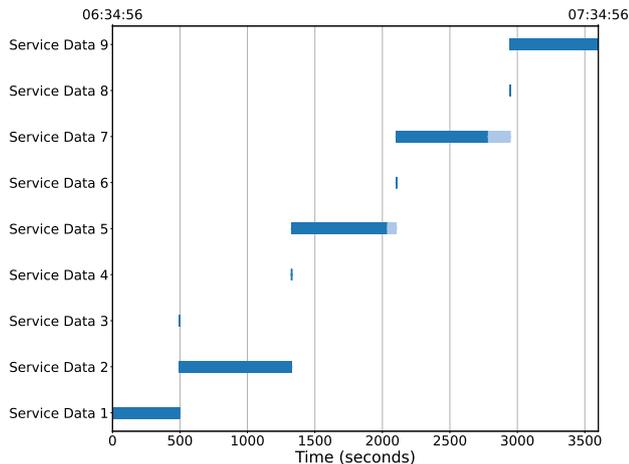


Fig. 6 Service Data changes in Google Pixel 6 Pro.

vice Data on AQUOS sense3 and ROG Phone 5s did not change. On Google Pixel 4a, the Service Data changed once during the experiment, while on Google Pixel 6 Pro, it changed five times. In contrast, packets with a Service Data length of 6 bytes were observed only on Google Pixel 4a and Google Pixel 6 Pro. These packets appeared roughly when the 27 byte Service Data changed and disappeared when the new 27-byte Service Data appeared. The frequency of Service Data changes on Google Pixel 6 Pro is shown in Fig. 6. Colors are assigned according to the order of BD addresses for each Service Data, and no identical BD address exists throughout the graph. The timing of Service Data switching indicates that these two types of packets are interlinked. However, in other experiments, the frequency of Service Data updates and their occurrence differed, and the conditions and rules governing packet transmission are not yet clarified.

4. BLE Advertising Packet-Based People Flow and Crowdedness Estimation in Real Environments

The authors have proposed several methods to estimate crowdedness levels (“BLECE”) and people flow (“BREID”) using BLE signals, especially BLE advertising packets containing randomized MAC addresses (random addresses), collected through fixed or mobile BLE scanners. These methods enable the estimation of both crowdedness and human movement patterns based on the collected BLE data.

In this section, we introduce a set of practical case studies applying these methods to various urban spaces and use cases. Through these examples, we explore the potential of BLE advertising packets as a versatile sensing modality for human mobility analysis, and discuss how such approaches can be leveraged.

4.1 BLECE: BLE-Based Crowdedness Estimation

The methods in this section aims to estimate how many people are present in or around the area where a BLE scanner is installed (i.e., crowdedness levels) using BLE signals.

Table 3 List of features used in the indoor crowdedness estimation model [4].

Feature Name ^{*a}	Description
all_num_ T sec	Total number of BD addresses scanned in the past T seconds
unique_num_ T sec	Number of unique BD addresses scanned in the past T seconds
unique_ratio_ T sec	Ratio of unique devices among the scanned BD addresses in the past T seconds (unique count/total count)
unique_num_ T sec_ S db	Number of unique BD addresses scanned in the past T seconds with RSSI greater than the threshold S dB
hour	Hour of the day (0–23)
is_weekday	Weekday: 1, Holiday: 0

^{*a} T represents the time range for referencing past samples (15, 30, 45, 60 seconds), and S denotes the threshold for referencing signal samples with high RSSI values (from -60 to -90 dB, in 5 dB increments).

Specifically, it targets scenarios such as counting people staying in indoor spaces like restaurants or public facilities, or estimating the number of passengers on public transportation such as buses or trains.

In this method, the random addresses contained in BLE advertising packets are used for filtering noise data, such as eliminating duplicate observations and excluding signals from individuals outside the area of interest. However, the core input for the crowdedness estimation model is data derived from the signal reception status.

The most basic set of input features used in this method is listed in Table 3, based on the features adopted in [4]. From the BLE advertising packets received by the scanner, various statistical data of BLE signals are calculated by changing the data acquisition time window (T), filtering for unique random addresses, and applying thresholds to RSSI (Received Signal Strength Indicator). These basic features are used as input data for machine learning to construct a model for estimating spatial crowdedness.

In addition to the basic features, incorporating location-specific features can further improve model performance. For instance, in the case of route buses [5], because passenger behavior varies by route, adding features such as route numbers and departure times from bus stops, or applying threshold conditions for the appearance rate of the same random MAC address within a target segment, can enhance estimation accuracy.

Practice 1: Crowdedness Estimation for Various Public Spaces

To capture the “dwell event” of people in a particular place in urban spaces, BLE scanners were installed in public spaces scattered throughout the city to estimate the number of people present in the target space.

In this practice [4], we targeted actual restaurants and public facilities, which differ in type, conditions, and size (four locations in total, as shown in Fig. 7). BLE data were collected for around 10 days using varying numbers of BLE scanners. The number of people in the space was obtained by counting manually or using CCTV cameras.

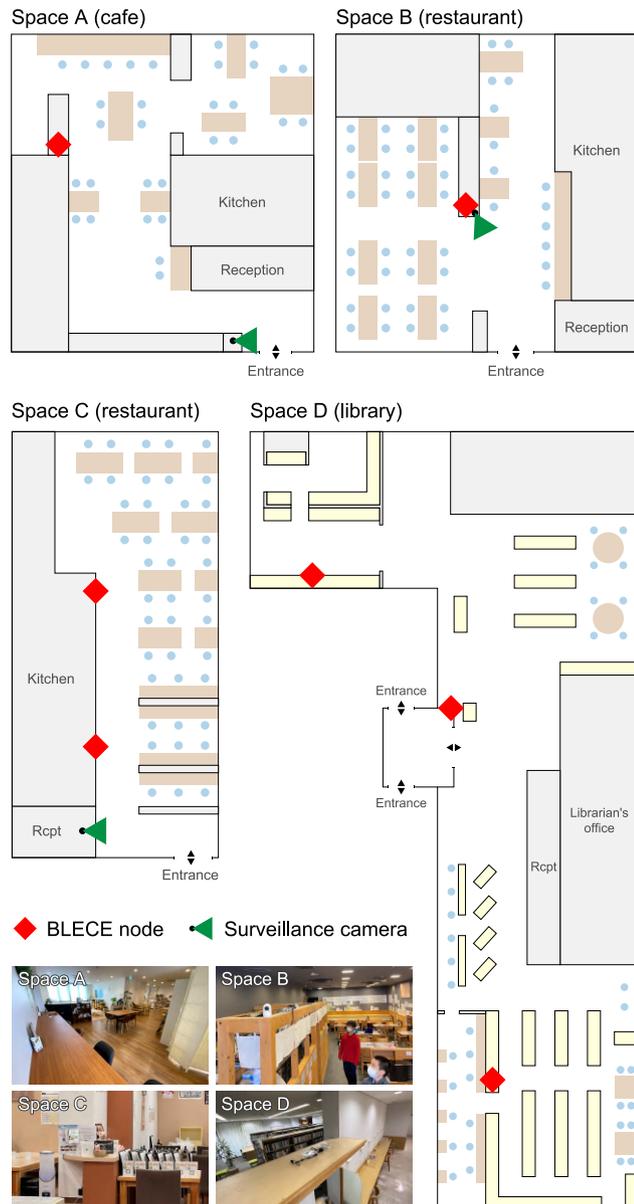


Fig. 7 Floor plan of experimental space in Practice 1 [4].

After data collection, we built and evaluated a crowdedness level estimation model. The results showed that the model built for each space using the XGBoost Regressor achieved a mean absolute error (MAE) of 4.89, a mean absolute percentage error (MAPE) of 84.0%, and a root mean square error (RMSE) of 6.34 in the worst case. These results indicate that a certain crowdedness level can be estimated using a model with common features.

Practice 2: Crowdedness Estimation for Public Transportation

The “dwell event” also occurs in spaces outside of buildings in urban areas. For example, in public transportation, people stay in the same space for a certain period, making it possible to estimate the number of people by installing BLE scanners

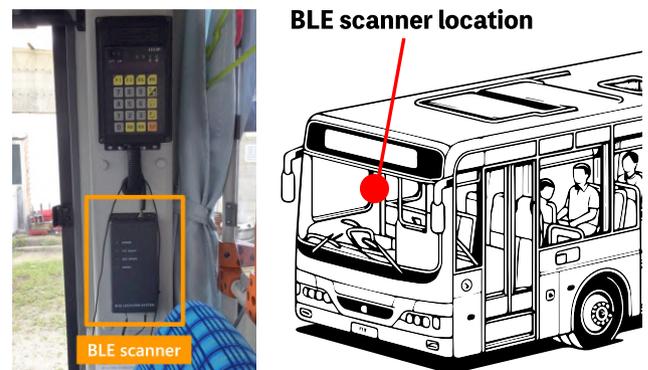


Fig. 8 BLE scanner installed to the existing onboard device of fixed-route bus, in Practice 2 [7].

inside the vehicles.

In the practical studies conducted by Kanamitsu et al. [5] and Taya et al. [6], the targets were actual route buses and railways operating in Nara Prefecture. BLE data were collected using BLE sensing devices built with Raspberry Pi, installed inside the vehicles. At the same time, the number of passengers for each travel segment was manually counted.

Using the manually counted passenger numbers as ground truth data and the BLE data as input data, a machine learning-based crowdedness estimation model was built and evaluated. As a result, the estimation for buses achieved a Mean Absolute Error (MAE) of 2.49 passengers and a Mean Absolute Percentage Error (MAPE) of 38.8%. For trains, the results were MAE of 5.56 passengers and MAPE of 27.0%.

Ikenaga et al. [7] further extended the research by Kanamitsu et al. by embedding BLE sensing functionality into the firmware of existing onboard devices (Fig. 8), enabling continuous BLE data collection. Based on the collected data, a model was constructed to estimate five levels of crowdedness (based on actual passenger count) and six levels of subjective crowdedness (how passengers perceive the degree of crowding). Evaluation results showed that the model could achieve 65.3% accuracy for actual crowdedness estimation (with a chance rate of 20.0%) and 51.3% accuracy for subjective crowdedness estimation (with a chance rate of 16.7%).

Practice 3: Crowdedness Estimation Using Mobile Scanners

A limitation of the BLE scanners used in the previous practices is that the measurable area is restricted to where the devices are physically installed. To address this, mobile devices are leveraged to act as BLE scanners, enabling crowdedness estimation at arbitrary locations.

In the practical study by Kikuchi et al. [8], a method was proposed that estimates crowdedness across entire urban areas by combining BLE scanning technology with participatory sensing. By using smartphones carried by users to scan nearby BLE signals, it becomes possible to collect data and estimate crowdedness over a wide area without installing dedicated devices. To examine the potential of the proposed method, an experiment was conducted in Nara

and Kyoto Prefectures, where BLE data from the surroundings, along with annotation data for the number of people nearby and perceived crowdedness (5 levels), were collected using smartphones. Machine learning models were trained to estimate both variables. As a result, the evaluation of the number of surrounding people showed a Mean Absolute Error (MAE) of 15.84, and the estimation of perceived crowdedness achieved an F1-Score of 0.29. Since the shape and layout of the target space were not provided in this evaluation, the difficulty of estimation task might be higher than other practices. By incorporating additional contextual information, the estimation accuracy could be improved in the future.

4.2 BReID: BLE-Device Re-Identification for People Flow Estimation

The methods in this section aims to understand how people move within areas where BLE scanners are installed, utilizing BLE signals to estimate human movement patterns. Specifically, this method targets the estimation of human trajectories in urban spaces as well as identifying boarding and alighting points in public transportation such as buses and trains.

This method tracks changes in random addresses contained in BLE advertising packets by utilizing an address carryover technique, enabling people flow estimation limited to the observation range of BLE scanners. Figure 9 illustrates the overview of address carryover when people move between scanners. When a random address changes while passing near scanners 1 and 2, the RSSI patterns in the observation data of each scanner form overlapping peaks. At a certain point in the time-series data, the previous address disappears, and a new address appears. The carryover algorithm links the pre-change and post-change addresses by leveraging hints such as the operating system of the transmitting device, the timing of address changes, and time-series RSSI variations.

When public transportation is targeted, the positional relationship between people and the onboard BLE scanner remains mostly unchanged during transit. Therefore, the address carryover algorithm determines the pair of addresses to be linked while considering this constraint. An example of address carryover result is shown in Fig. 10. Black dotted arrow shows timing when the address carryover algorithm applies.

There are several candidate types of packets that can be used. Kawashima et al. [11] proposed a method that focused on specific packets transmitted by the COCOA (COVID-19 Contact-Confirming Application)[†], which was used during the COVID-19 pandemic. However, since the service was discontinued in November 2022, these packets are no longer available.

In contrast, based on the findings from the packet transmission investigation described in Sect. 3, it is possible to

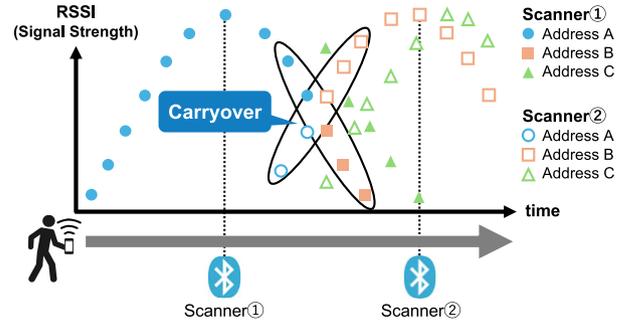


Fig. 9 Address carryover mechanism for BLE advertising packets in Practice 4 [33].

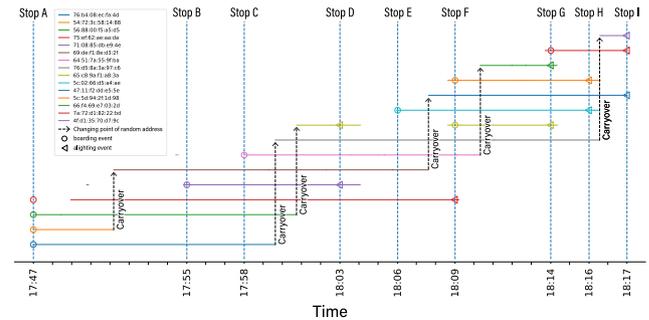


Fig. 10 Address carryover result example for origin-destination estimation of fixed-route bus in Practice 5 [34].

achieve similar people flow estimation by targeting specific types of packets: those beginning with `0x4c0010` in the Manufacturer Specific Data (MSD) field for iOS devices, and those including the Service Data - 16-bit UUID `0xfeF3` for Android devices.

Practice 4: People Flow Estimation for City

To capture the movement of people in urban spaces, multiple BLE scanners were installed throughout the city to estimate movement between scanners.

In the practice conducted by Hayashi et al. [9], 30 BLE scanners (based on Raspberry Pi 4) were installed over a 2.5 km area centered around Saga Station in Saga City, Saga Prefecture (Fig. 11). The experiment was conducted over 60 days, during which specific BLE packets were collected based on the investigation described in Sect. 3. During the experiment period, GPS logs were collected and used as ground truth trajectory data to evaluate the proposed method.

As a result, the average coverage of the estimated trajectory data obtained by the proposed method for the ground truth trajectory data was 70.9% for walking as a means of transportation, 46.2% for cycling, and 66.0% for all cases. By applying the address carryover technique, the average coverage improved compared to the baseline method as follows: 9.5 percentage points for walking, 3.4 percentage points for cycling, and 8.3 percentage points for overall result. These results demonstrate the effectiveness of the address carryover technique in enhancing the accuracy of people flow estimation, especially when the movement pattern is

[†] <https://www.digital.go.jp/policies/cocoa>

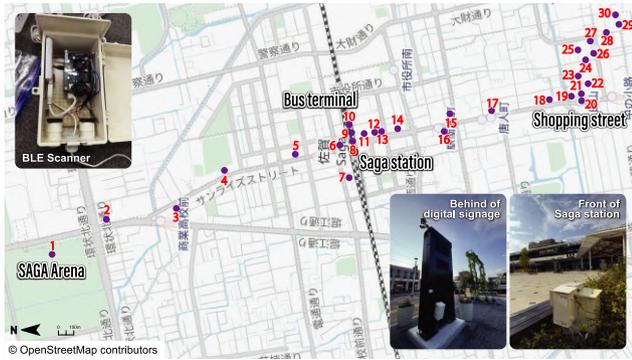


Fig. 11 BLE scanner location in Practice 4 [33].

predominantly walking. The proposed method significantly outperforms the baseline in capturing human mobility in urban environments.

In the above practical experiment, the estimation of movement trajectories did not take into account the direction of human movement. However, providing information about the movement direction has the potential to improve the accuracy of trajectory estimation. Goto et al. [10] proposed a method for estimating the movement direction of people in places where there is significant pedestrian traffic, such as streets and shopping arcades. This method primarily uses the “time-series difference in RSSI” calculated from the data obtained by two BLE scanners installed along the street. The evaluation results showed that in a crowded street in Midosuji at Osaka Prefecture with approximately 80 people passing by on average, this method could estimate the movement direction with a mean absolute error of 18.6 and 16.0 for each direction, respectively.

Practice 5: Origin-Destination Estimation for Public Transportation

To capture the movement of people using public transportation, such as route buses, BLE scanners were installed inside buses to estimate the origin and destination (OD) data — namely, where passengers board and alight.

In the practical experiment conducted by Kawashima et al. [11], the target was an actual route bus operating in Kyoto City, Kyoto Prefecture (Fig. 12). Data collection involved adding BLE scanning functionality to the digital tachograph device of the route bus to collect BLE data transmitted from COCOA (COVID-19 Contact-Confirming Application) while simultaneously manually recording the OD data of each passenger. As a result, the precision ranged from 60% to 100% when considering the bypassed bus stops. However, the recall was relatively low — less than 20% on average — primarily due to the limited data share from COCOA.

In the practical experiment conducted by Yoshimura et al. [12], the target was an actual route bus operating in Okayama City, Okayama Prefecture. Data collection was performed using BLE scanners based on Raspberry Pi 4, the same setup as used by Hayashi et al. [9]. Specific BLE pack-



Fig. 12 BLE scanner installed to the existing digital tachograph device in Practice 5 [11].

ets were collected based on the investigation described in Sect. 3. As a result, the OD estimation achieved an average precision of 70% (max: 100%, min: 53%) and an average recall of 34% (max: 59%, min: 15%). This demonstrates that the proposed method maintains a level of precision comparable to the use of COCOA packets, while significantly improving the recall rate.

5. Conclusion

In this study, we aim to establish a method for estimating the number of people staying in urban spaces (crowdedness) and their movement between locations (people flow) by utilizing data from BLE advertising packets — such as RSSI and randomized addresses — transmitted by personal smartphones and other devices. This paper presented the results of a preliminary investigation on the characteristics of BLE advertising packets, an overview of the proposed estimation methods for crowdedness level and people flow based on those findings, and five practical field experiments conducted in real-world environments.

As future work, we plan to integrate the individual estimation methods for crowdedness and people flow to develop a unified approach for tracking human behavior in terms of both “dwell” and “movement.” Furthermore, we aim to establish a comprehensive method for understanding crowdedness and people flow across an entire urban environment by proposing techniques that can capture human mobility across different types of spaces — such as public spaces and public transportations.

Acknowledgments

This study was supported in part by JST PRESTO under Grant No. JPMJPR2039 and JST COI-NEXT under Grant No. JPMJPF2115. We would like to express our deep gratitude to everyone at Ikoma City, Osaka City, Saga City, Nara Kotsu Co., Ltd., Kyoto Bus Inc., Chutetsu Bus Co., Ltd., Kintetsu Railway Co., Ltd., Local Media Labs K.K, and Restaurants in Nara for their significant support in conducting this research.

References

- [1] Ministry of Land, Infrastructure, Transport and Tourism, Japan, "Guidelines for Utilizing People Flow Data to Solve Regional Issues Ver1.2," https://www.mlit.go.jp/tochijudousan_ensetsugyo/chirikukannjoho/tochijudousan_ensetsugyo_k170000100034.html, 2024. Accessed on March 21, 2025 (in Japanese).
- [2] M. Versichele, T. Neutens, M. Delafontaine, and N. Van de Weghe, "The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities," *Applied Geography*, vol.32, no.2, pp.208–220, 2012.
- [3] M. Mochizuki, T. Onikura, Y. Fukuzaki, and N. Nishio, "A pedestrian flow analysis system using Wi-Fi packet sensors and its application to a real environment," *Proc. Multimedia, Distributed, Cooperative and Mobile Symposium (DICOMO'14)*, pp.1249–1257, 2014.
- [4] Y. Matsuda, K. Ueda, E. Taya, H. Suwa, and K. Yasumoto, "BLECE: BLE-based crowdedness estimation method for restaurants and public facilities," *14th International Conference on Mobile Computing and Ubiquitous Networking (ICMU'23)*, pp.1–6, 2023.
- [5] Y. Kanamitsu, E. Taya, K. Tachibana, Y. Nakamura, Y. Matsuda, H. Suwa, and K. Yasumoto, "Estimating congestion in a fixed-route bus by using ble signals," *Sensors*, vol.22, no.3, pp.1–15, 2022.
- [6] E. Taya, Y. Kanamitsu, K. Tachibana, Y. Nakamura, Y. Matsuda, H. Suwa, and K. Yasumoto, "Estimating congestion in train cars by using ble signals," *2nd Workshop on Data-Driven and Intelligent Cyber-Physical Systems for Smart Cities (DI-CPS'22)*, pp.1–7, 2022.
- [7] T. Ikenaga, Y. Matsuda, I. Goto, K. Ueda, H. Suwa, and K. Yasumoto, "Using BLE signals to estimate objective and subjective crowdedness levels on fixed-route buses," *IEEE Access*, vol.13, pp.67488–67499, 2025.
- [8] T. Kikuchi, T. Ikenaga, R. Hosokawa, Y. Matsuda, H. Suwa, and K. Yasumoto, "Exploring urban crowding estimation through participatory BLE scanning," *Proc. 2024 Kansai Branch Convention of IPSJ*, 6 pages, 2024.
- [9] K. Hayashi, T. Yoshimura, H. Kobayashi, R. Imamura, M. Itou, K. Ushijima, H. Suwa, and Y. Matsuda, "Case study of people flow estimation in urban space using BLE advertising packets ~Data analysis and discussion during the period of the JAPAN GAMES in Saga City~,," *IPSI SIG Technical Report*, 2025.
- [10] I. Goto, K. Ueda, Y. Matsuda, H. Suwa, and K. Yasumoto, "BLESS: BLE based street sensing for people counting and flow direction estimation," *2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom'24)*, pp.76–81, 2024.
- [11] M. Kawashima, I. Arai, A. Endo, M. Kakiuchi, and K. Fujikawa, "Origin destination estimation carrying over rolling proximity identifiers with RSSI," *2022 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT)*, pp.7–12, 2022.
- [12] T. Yoshimura, K. Hayashi, I. Arai, and Y. Matsuda, "Origin-destination estimation for fixed-route bus using BLE advertising packets ~Data analysis and discussion through experiments on two actual operating routes~,," *IPSI SIG Technical Report*, 2025.
- [13] NTT DOCOMO, INC., "Mobile Spatial Statistics," <https://mobaku.jp/>, (Accessed on March 21, 2025), 2008.
- [14] AGOOP Corp., "Kompreno," <https://www.agoop.co.jp/service/kompreno/>, (Accessed on March 21, 2025), 2018.
- [15] V.A. Sindagi and V.M. Patel, "Generating high-quality crowd density maps using contextual pyramid CNNs," *Proc. IEEE International Conference on Computer Vision*, pp.1861–1870, 2017.
- [16] W. Liu, M. Salzmann, and P. Fua, "Context-aware crowd counting," *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR'19*, pp.5099–5108, 2019.
- [17] SICK AG, "Counting people with 3D-LiDAR sensors," <https://www.sick.com/us/en/industries/building-management/access-control-gates/counting-people-with-3d-lidar-sensors/c/p616144>, (Accessed on 03.14.2023), 2020.
- [18] P.G. Kannan, S.P. Venkatagiri, M.C. Chan, A.L. Ananda, and L.S. Peh, "Low cost crowd counting using audio tones," *Proc. 10th ACM Conference on Embedded Network Sensor Systems, SenSys'12*, pp.155–168, ACM, 2012.
- [19] T. Nishimura, T. Higuchi, H. Yamaguchi, and T. Higashino, "Detecting smoothness of pedestrian flows by participatory sensing with mobile phones," *Proc. 2014 ACM International Symposium on Wearable Computers, ISWC'14*, pp.15–18, ACM, 2014.
- [20] M. Elhamshary, M. Youssef, A. Uchiyama, A. Hiromori, H. Yamaguchi, and T. Higashino, "CrowdMeter: Gauging congestion level in railway stations using smartphones," *Pervasive and Mobile Computing*, vol.58, p.101014, 2019.
- [21] H. Song, X. Liu, X. Zhang, and J. Hu, "Real-time monitoring for crowd counting using video surveillance and GIS," *2012 2nd International Conference on Remote Sensing, Environment and Transportation Engineering*, pp.1–4, IEEE, 2012.
- [22] E.U. Haq, X. Huarong, C. Xuhui, Z. Wanqing, F. Jianping, and F. Abid, "A fast hybrid computer vision technique for real-time embedded bus passenger flow calculation through camera," *Multimed. Tools Appl.*, vol.79, no.1–2, pp.1007–1036, 2020.
- [23] M. Handte, S. Foell, S. Wagner, G. Kortuem, and P.J. Marrón, "An internet-of-things enabled connected navigation system for urban bus riders," *IEEE Internet Things J.*, vol.3, no.5, pp.735–744, 2016.
- [24] A. Hidayat, S. Terabe, and H. Yaginuma, "Estimating bus passenger volume based on a Wi-Fi scanner survey," *Transportation Research Interdisciplinary Perspectives*, vol.6, p.100142, 2020.
- [25] Y. Maekawa, A. Uchiyama, H. Yamaguchi, and T. Higashino, "Car-level congestion and position estimation for railway trips using mobile phones," *Proc. 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp'14*, pp.939–950, ACM, 2014.
- [26] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," *2013 IEEE International Conference on Pervasive Computing and Communications, PerCom'13*, pp.193–200, ACM, 2013.
- [27] R. Takahashi, K. Hayashi, Y. Mitsukude, M. Futamata, S. Inoue, S. Matsuo, S. Ishida, Y. Arakawa, and S. Takano, "Itocon — A system for visualizing the congestion of bus stops around ito campus in real-time," *18th ACM Conference on Embedded Networked Sensor Systems, SenSys'20*, pp.697–698, 2020.
- [28] G. Wang, J. Lai, P. Huang, and X. Xie, "Spatial-temporal person re-identification," *Proc. AAAI Conference on Artificial Intelligence*, pp.8933–8940, 2019.
- [29] Y. Nagata, K. Hiroi, T. Yonezawa, and N. Kawaguchi, "Privacy preserving person-flow estimation using multiple 3D people counter," *Multimedia, Distributed, Cooperative, and Mobile Symposium (DICOMO'20)*, pp.1020–1028, 2020.
- [30] H. Yamaguchi, A. Hiromori, and T. Higashino, "A human tracking and sensing platform for enabling smart city applications," *Proc. Workshop Program of the 19th International Conference on Distributed Computing and Networking, Workshops ICDCN'18*, pp.1–6, 2018.
- [31] M. Versichele, T. Neutens, M. Delafontaine, and N. Van de Weghe, "The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities," *Applied Geography*, vol.32, no.2, pp.208–220, 2012.
- [32] D. Murai, K. Urano, M. Mochizuki, T. Yonezawa, J. Nishida, and N. Kawaguchi, "Influence on Wi-Fi packet human flow sensor in large-scale outdoor facilities and verification of its utilization," *IPSI Trans. Digital Practices*, pp.12–21, 2022.
- [33] K. Hayashi, T. Yoshimura, H. Suwa, and Y. Matsuda, "Exploration of city-wise people flow estimation method using multiple BLE scanners," *Proc. 2024 Kansai Branch Convention of IPSJ*, pp.1–8, 2024.
- [34] T. Yoshimura, K. Hayashi, I. Arai, and Y. Matsuda, "Initial study on OD data estimation using BLE random MAC addresses," *Proc. 2024 Kansai Branch Convention of IPSJ*, pp.1–8, 2024.
- [35] Bluetooth SIG, "Assigned Numbers," Nov. 2024. Version Date:

2024-11-06. Available: <https://www.bluetooth.com/specifications/assigned-numbers/>



Yuki Matsuda received the B.E. degree in advanced course of mechanical and electronic system engineering from the National Institute of Technology, Akashi College, Japan, in 2015, and the M.E. and Ph.D. degrees from the Graduate School of Information Science, Nara Institute of Science and Technology, Japan, in 2016 and 2019, respectively. He is currently a Lecturer at Faculty of Environmental, Life, Natural Science and Technology, Okayama University, Japan. His current research interests include participatory sensing, location-based information systems, wearable computing, and affective computing. He is a member of IPSJ, IEICE, JSAI, ACM, and IEEE.

participatory sensing, location-based information systems, wearable computing, and affective computing. He is a member of IPSJ, IEICE, JSAI, ACM, and IEEE.



Hirohiko Suwa received the B.E. and M.E. degrees from Gunma University, in 1998, and the Ph.D. degree from the University of Electro-Communications, Japan, in 2000 and 2006, respectively. He was an Assistant Professor with the University of Electro-Communications, from 2006 to 2014, and the Nara Institute of Science and Technology (NAIST), from 2014 to 2018, where he has been an Associate Professor, since 2019. His research interests include social informatics, real data mining, and ubiquitous computing systems. He is a member of IEEE, IEICE and IPSJ.

computing systems. He is a member of IEEE, IEICE and IPSJ.



Kotaro Hayashi received the B.E. degree from Okayama University in 2025. He is currently a graduate student at Graduate School of Environmental, Life, Natural Science and Technology, Okayama University. His research interests include sensor network, urban sensing, people flow analysis. He is a member of IPSJ.



Taito Yoshimura received the B.E. degree from Okayama University in 2025. He is currently a graduate student at Graduate School of Environmental, Life, Natural Science and Technology, Okayama University. His research interests include intelligent transport systems, urban sensing, people flow analysis. He is a member of IPSJ.



Arata Yoshihara received the B.E. degree in Information Sciences from Oita University in 2023 and the M.E. degree from the Graduate School of Science and Technology, Nara Institute of Science and Technology, in 2025. His research interests include wireless communication technologies.



Ismail Arai received the M.E. and Ph.D. degrees in engineering from the Graduate School of Information Science, Nara Institute of Science and Technology, in 2004 and 2008, respectively. From 2008 to 2011, he was a Postdoctoral Fellow with the Research Organization of Science and Engineering, Ritsumeikan University. He was an Assistant Professor, a Lecturer, and an Associate Professor with the Department of Electrical and Computer Engineering, National Institute of Technology, Akashi College, from 2011 to 2013,

from 2013 to 2015, from 2015 to 2016, respectively. He has been an Associate Professor affiliated with the Information Initiative Center, Nara Institute of Science and Technology, since 2016. His research topics include transportation data analysis, pedestrian navigation systems, cybersecurity, and open data. He is also a member of the ACM, IEEE, and IPSJ.