On the Relationship Between Accuracy of Bus Position Estimated by Crowdsourcing and Participation Density

Kenro Aihara[1,2,3(✉)], Piao Bin[1], and Hajime Imura[4]

[1] National Institute of Informatics,
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
kenro.aihara@nii.ac.jp
[2] The Graduate University for Advanced Studies, Hayama, Japan
[3] Joint Support-Center for Data Science Research,
Research Organization of Information and Systems, Tokyo, Japan
[4] Hokkaido University, N-13, W-8, Sapporo, Hokkaido 060-8628, Japan

Abstract. The authors proposed a methodology of bus location service by crowdsource. In the conventional bus location service, a GPS receiver or the like for positioning the bus vehicle and a communication line for transmitting the position information are necessary for each bus and it causes expensive cost, whereas in this methodology, the BLE beacon and the application of the service installed in users' smartphone transmits the beacon signal and the position information of the smartphone using the communication line of users' smartphone so as to reduce the cost of bus operators.

In order to grasp the position of each bus at any time on the server, one of the following is required: (1) At least one service user always exists for each bus, and (2) It is possible to estimate the position with high accuracy where there is no service user in a section.

If contributing to the realization of smart city by grasping the situation in the town by crowdsourcing, it is extremely important to clarify the precision with respect to density that may be given as a relation between the number of users who participate in the program for grasping this situation and the size of the town.

In this paper, the authors examine and show the relationship between the density of data collection and the accuracy of interpolation for the section where data is missing, using the position information actually obtained using the proposed application.

Keywords: Internet of Things · Smart and hybrid cities · Crowdsensing · Crowdsourcing

1   Introduction

We have researched and proposed crowdsourcing which can collect and share the location information of mobile phones cheaply in a public place like a bus. Citizen participation is an indispensable factor in the realization of smart city, but crowdsourcing, which enables this technically, is an important issue. In order to grasp the situation of the city by crowdsourcing, the density of data collection is necessary. For example, if you want to grasp the traffic volume of a certain area at any time, you need to collect the traffic conditions of various places in the area by crowdsourcing, but for some partial areas only once in several days It is not always said that it is insufficient to grasp traffic conditions of partial areas. The required density depends on the quality of the service, but if you know the amount of traffic per hour for the entire area, for example when it is defined in the service, it is simply at least one question per hour. It is necessary to provide information.

Here, as an approach to realize a service with a lower density, it is conceivable to accurately interpolate or estimate a portion where data is missing. In the case of the traffic volume on the road, since the constraints such as the network topology can be used, the density in a planar sense is not necessary, and the state of the unknown section may be estimated from the information of the surrounding road sections to be connected. Therefore, to maintain the service quality, it neccessary to obtain the balance between the estimation accuracy and what degree of density of collected data can be obtained.

In this paper, we will examine the estimation accuracy of the density of the user and its surrounding "data collector" and the bus position by simulation based on actually acquired data by taking the location information service of the bus by crowdsourcing as an example.

2   Background

2.1   Crowdsourcing for Civil Problems

The term "crowdsourcing" was described by Howe in 2006 [6] and defined that crowdsourcing is the act of taking a task traditionally performed by a designated agent and outsourcing it by making an open call to an undefined but large group of people [7]. This can take the form of peer-production, but is also often undertaken by sole individuals [5].

The concept of smart cities can be viewed as a recognition of the growing importance of digital technologies for a competitive position and a sustainable future [10]. Although the smart city-agenda, which grants ICTs with the task to achieve strategic urban development goals such as improving the life quality of its citizens and creating sustainable growth, has gained a lot of momentum in recent years.

Tools such as smartphones offer the opportunity to facilitate co-creation between citizens and authority. Such tools have the potential to organize and stimulate communication between citizens and authority, and allow citizens to

participate in the public domain [4,12]. One example is FixMyStreet[1] that enables citizens to report broken streetlights and potholes [9]. It is important that these approaches will not succeed automatically and social standards like trust, openness, and consideration of mutual interests have to be guaranteed to make citizen engaging in the public domain challenging.

Waze[2] is another crowdsourcing service to collect data of traffic. Even though Waze provides users to traffic information collected from users and route navigation function, it seems not enough to motivate users to get involved in, because recommended routes are not as adequate as car navigation appliances, especially in Japan where such appliances are well-developed.

2.2   Bus Location Services

Route bus system is a fundamental transit service. However, due to a progress of motorization, the number of bus passengers has been gradually decreasing especially in suburban areas. As a result, the decline in passengers has led to a decline in unprofitable routes and it is in a vicious circle that accelerates the decline of passengers. To attract more choice passengers to route buses, the transit service must not only have a high level of service in terms of frequency and travel time but also must be reliable [13]. Although such efforts often come at a substantial cost, one inexpensive way to improve unreliability from the user perspective is providing real-time transit information.

A bus location service, or realtime bus tracking service, provides up-to-date bus location and estimated times of arrival at bus stops. Most existing bus location systems (e.g., [8,11]) use onboard location sensors, such as GPS receivers, to perceive the current location and then send it to a server. In these systems, it is necessary to prepare a communication line, such as cellular phone network, and GPS receiver in advance.

Another model is a bus detector installed on the environment side along the route detects the ID of the nearby bus and sends it to the server. Typically, these detectors are installed at the bus stop and detect the passing bus. Detectors installed at the bus stop need a communication line for transmitting data to the server.

Both models have high initial cost of equipment and operational cost of communication, and there is a problem in introducing services.

3   Target Service: Ride Around-the-Corner

Here we describe the target service of crowdsourced bus location service "Ride around-the-corner. (Ride ATC)" in blief [2]. The key idea of the service model is that collecting bus locations is not by bus operators but by onboard passengers. To collect them, a smartphone application of bus tracker is provided to public.

---

[1] https://www.fixmystreet.com/.
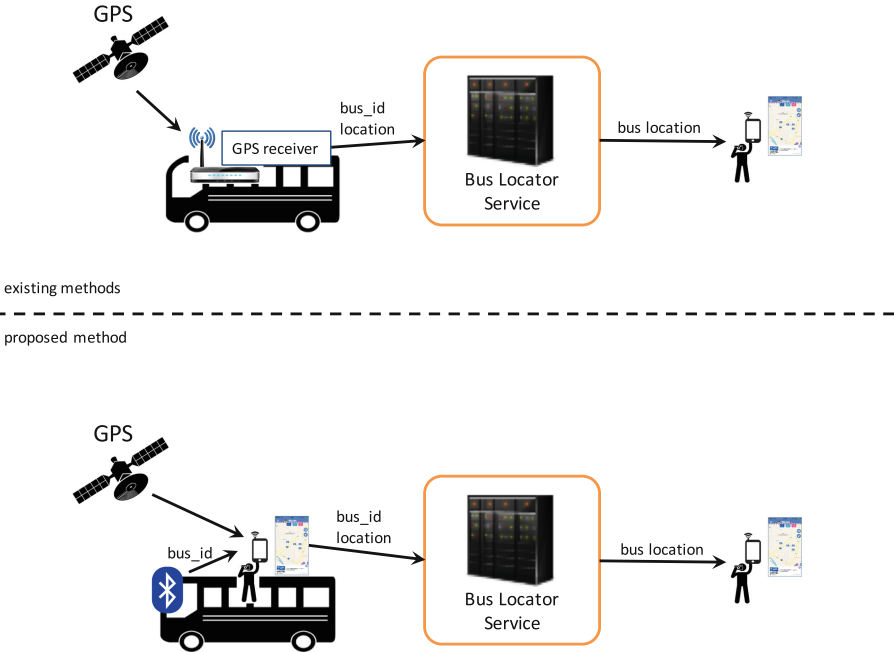[2] https://www.waze.com/.

Fig. 1. Comparison of methods

Figure 1 shows the proposed method, compared with existing ones. Typical existing methods are equipped with onboard location sensors, such as a GPS receivers, and transmitters. Every perceived latest location will be sent to the server frequently through a wireless communication line, such as cellular phone. That is, bus operators must prepare such devices and a communication line for each bus vehicle.

On the other hand, in this model, bus operators only install Bluetooth beacons on each bus vehicle. Instead of installing a location sensor on the bus, the proposed method uses a passenger's smartphone and our application. An onboard beacon broadcasts its own identification number nearby. In the case of Bluetooth, its range is usually several tens of meters. When the application on passenger's smartphone detects specific Bluetooth information including UUID and vehicle identification number, it perceives its location and transmit it with the vehicle information to the server.

The application shows current locations of buses in operation on bus transit services, while it detects nearby buses around users and transmits bus IDs with time and location of detection to the service platform. That is, locations of buses are collected by users.

3.1   Onboard Beacon

For enabling mobile applications detect IDs of buses, Bluetooth beacons are deployed on buses. In our preliminary experiment, we set one beacon for each bus vehicle, because the range of Bluetooth signal is usually several tens of meters and that can cover the whole vehicle.

Each beacon broadcasts the common service UUID and its own identification number. For IDs, we use the same major number for bus vehicles and an unique minor number. Ids of buses and relations to the fleet are stored in a database in the server. In addition to IDs of buses, some static information, such as bus routes and timetables, are given to the server database. The system can identify which bus on which route the beacon is on by the correspondence table of major and minor number, vehicle and route in the database.

3.2   Mobile Applications

We are preparing two as a mobile application related to this Ride ATC. One, of course, is the Ride ATC application itself. It is an application that can check the bus position used by the user wishing to know the bus position. The other is "Drive around-the-corner. (DriveATC)" application which is a related service.

These two applications differ in target users, but both have the function of notifying the bus position information when detecting the bus beacon signal. Therefore, the location information of the bus will be collected from various mobiles who act in the streets. The idea of around-the-corner, crowdsourcing service is to promote understanding of the situation of the city as a whole by sharing necessary information mutually between services in this manner.

Ride ATC Application. The Ride ATC application provides users with current locations of buses in operation. It indicates the position of the buses on the map. The user can check it and estimate arrival time of their possible target bus on map beforehand so that it is possible to minimize the waiting time at the bus stop without missing it.

While the user is using the application in both the foreground and the background, the Ride ATC application scans a specific range of Bluetooth IDs. Once it detects an ID nearby, the application obtains location from onboard sensors. The location data with bus IDs and time are collected and pooled in the local data store and then transmitted to the service platform.

On the platform, current positions are updated and estimated by using collected bus IDs with spatiotemporal information and static information in database.

Drive ATC Application. The authors have developed and provides a driving recorder service called "Drive around-the-corner." since February 2015 [1,3]. Drive ATC has the function of collecting behavior logs and posts of events and delivering information around current position.

Before driving, users mount their own smartphone and connect a cable for power supply if neccessary (Fig. 2), and then start recording in the application (Fig. 3).

Drive ATC users naturally use this application as a drive recorder, but when the application simultaneously detects a beacon signal of the neighboring ATC service, such as onboard beacons on the bus, it has the function of sending it to the server together with the position information like the Ride ATC application.
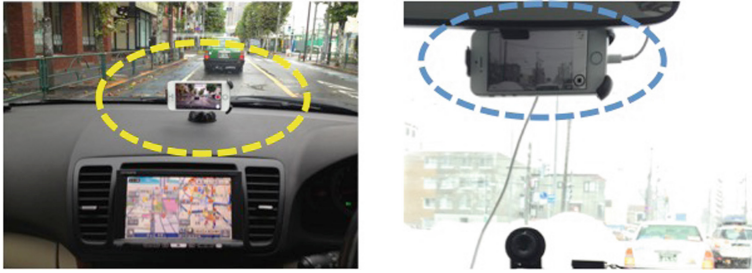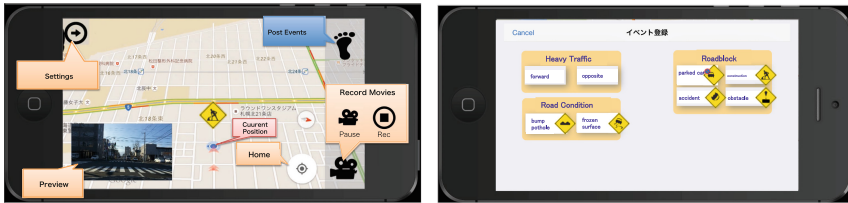

Fig. 2. Mounting smartphones on car.

(a) main screen   (b) post events

Fig. 3. "Drive: around-the-corner." application. Traffic information, events posted from users, events extracted from sensor data, and footprints are shown in map of main screen.

3.3   Grasping Bus Location and Estimating Arrival Time

If at least one passenger who enables the Ride ATC is onboard, the service grasps the current location of the bus. Therefore, at least one user of the Ride ATC gets on a bus, this service gets effective. Conversely, it is not possible to acquire the location information of the bus where there is no user onboard. In other words, the more users want to know the location of the bus, the more practical this service will be.

The signal of the beacon on the bus can be received not only within the bus but also in the vicinity of the outside of the bus (Fig. 4). For example, in the case where the bus passes by the side of the user walking on the sidewalk, the user can provide the location information of the bus to the service.
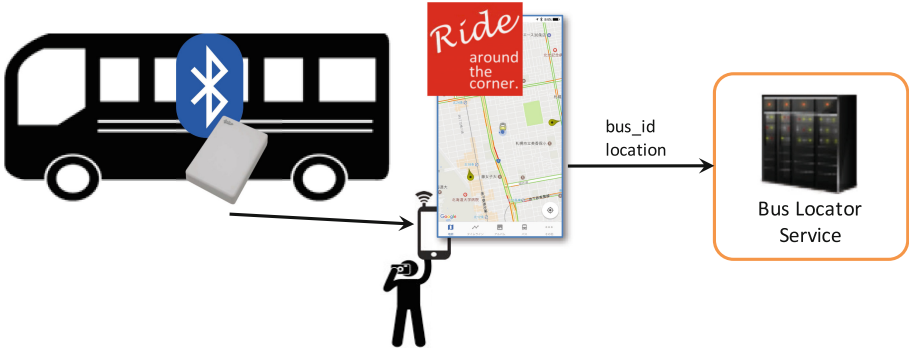
Fig. 4. Detecting onboard beacons from outside the bus

Table 1. Collected data of Ride ATC

Detector	Number of records
By onboard passengers	12,346
From outside of the bus	1,920

4   Density and Accuracy

4.1   Accuracy

The accuracy of position information is discussed from two viewpoints.

On the other hand, regarding the density of data, it is considered that positioning and communication are carried out generally once per minute in the conventional method of installing a GPS receiver on the bus. In the method of installing the detector at the bus stop, since the position of the bus is detected only when the bus approaches the bus stop, its density is smaller and it is impossible to grasp the behavior between bus stops. In the proposed method, the user can continuously measure and transmit the position information while detecting the beacon, and in the current implementation of the Ride ATC, these are performed every second.

4.2   Acquired Data

We conducted experiments on data acquisition in cooperation with the Hokkaido Chuo Bus in March 2018. We prepared passengers who are eight users of the application, and acquired data over specific routes over two days. We set up a beacon on 38 buses used for six routes and set up a beacon at 150 bus stops on that route. Table 1 shows number of data collected at the experiments. This table shows the number of beacons detected by passengers over two days of use.

When getting on the bus where the beacon was installed, the beacon is detected without problems and sent to the server. Although detailed verifica-

tion was not completed, beacon detection was able to be performed in the range of roughly 50 m.

What the authors want to pay attention to here is that it can detect the bus beacon even from outside the bus. Here, 1920 out of all 14,266 detections was other than the bus that passengers themselves take. These include the case of detecting a beacon of another bus passed by both while getting on another bus and while on the roadside. This indicates that this service can detect the location of the bus even if it is not necessarily on the target bus. The authors suppose that users of Drive ATC can also detect the bus beacon and provide information.

4.3   Settings

Service Quality. As a service quality, here we define an interval to update the bus location is 15 s as a baseline. One reason is that the traditional bus location services in the urban bus route where the bus stops are short and then the information is too coarse at intervals longer than 1 min. And also in the major bus location service operated in Japan, such as Business Navitime Bus Location Service, which collects the location information by installing the device onboard, the update frequency is about 15 s.

Evaluation Index. In the method of crowdsourcing targeted here, since it depends on whether or not the user is on the bus, this paper regards each section (between bus stops) as a unit. In normal operation, in addition to business related factors, such as existence of passengers who get on or off at the bus stop, it is impossible to move in the scheduled travel time due to external factors such as traffic signals, road congestion and obstacles on the road. Here, the influence on these estimation accuracy for each section s at i^{th} run is defined as an error rate as follows.

$$r_{si} = \frac{(observed\ travel\ time)_i}{(expected\ travel\ time)_i} \tag{1}$$

This error rate indicates how deviated per 1 min travel time.

Data Blank. As a result of the experiment, it seems that location is detected almost every second inside the range of beacon signal, although the beacon detection timing depends on the OS of the smartphone. Therefore, if at least one user is aboard, this Ride ATC notifies the location of the bus which changes moment every second, and it can be said that it realizes real time positioning superior to the baseline service notified every 15 s. Figure 5 shows an example of collected data. The red triangle marker indicates the position and movement direction. The size of the marker depends on the speed. Naturally, the moving speed is changing every moment.

Fig. 5. An example of collected data on map

Target Route. In this paper, the authors focus on the bus route passing through the area around Sapporo terminal station where traffic congestion occurs burstingly in Sapporo city and the travel time of the bus is greatly affected because the necessity of bus location service is low in the case of routes whose schedule is not much disturbed.

Using the actual movement record of this route gathered using Ride ATC, the error rate was calculated from the running condition of each route section.

We selected a bus route of Hokkaido Chuo Bus between Sapporo terminal station and Asabu bus terminal. The route consists of twelve bus stops and the total journey is about 4.3 km. Usually it takes 18 min. However, if there is snowfall in the winter season, it may take twice this time in severe cases.
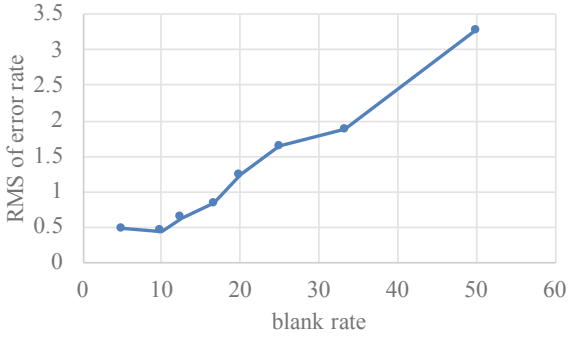
Fig. 6. Error rate

Table 2. Relationship between error rate and data blank rate

Blank rate (%)	RMS of error rate
5.0	0.487
10.0	0.443
12.5	0.635
20.0	1.237
25.0	1.641
50.0	3.269

4.4   Results

The root mean square of the error rate calculated from the actual data acquired in this route was 0.866. Among the actually acquired data, there was a delay of three minutes for the scheduled route of 18 min. Also, there was one that took four minutes in the section where the travel time was one minute on the timetable.

Figure 6 and Table 2 show a result of the relationship between error rate and data blank rate. The data blank section is randomly extracted at a given ratio, the value of root mean square of the error rate is calculated, and the average value of five tests is shown.

As a result, if the blank rate is less than this, it can not be said that it can be used as a service, but it can be said that it can maintain a certain level of service even for intermittent data collection on an congested route. It is suggestive that the blank rate at the same level as actual RMS, 0.866, was just below 20%.

5   Conclusion

In this paper, the authors examine and show the relationship between the density of data collection and the accuracy of interpolation for the section where data is missing, using the position information actually obtained using the proposed application.

The authors have been doing preliminary experiments. Under the cooperation of Sapporo city and the Hokkaido Chuo Bus, we are proceeding with the operation verification using the developed Ride ATC application. We will proceed with verification of methodology and social implementation by carrying out larger-scale demonstration.

References

1. Aihara, K., Bin, P., Imura, H., Takasu, A., Tanaka, Y.: On feasibility of crowdsourced mobile sensing for smarter city life. In: Streitz, N., Markopoulos, P. (eds.) DAPI 2016. LNCS, vol. 9749, pp. 395–404. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39862-4_36

2. Aihara, K., Bin, P., Imura, H., Takasu, A., Tanaka, Y.: Collecting bus locations by users: a crowdsourcing model to estimate operation status of bus transit service. In: Streitz, N., Konomi, S. (eds.) DAPI 2018. LNCS, vol. 10921, pp. 171–180. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91125-0_14

3. Aihara, K., Imura, H., Piao, B., Takasu, A., Tanaka, Y.: Mobile crowdsensing to collect road conditions and events. In: Yasuura, H., Kyung, C.-M., Liu, Y., Lin, Y.-L. (eds.) Smart Sensors at the IoT Frontier, pp. 271–297. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-55345-0_11

4. Amichai-Hamburger, Y.: Potential and promise of online volunteering. Comput. Hum. Behav. **24**(2), 544–562 (2008)

5. Howe, J.: Crowdsourcing: a definition. Crowdsourcing tracking the rise of the amateur (2006)

6. Howe, J.: The rise of crowdsourcing. Wired Mag. **14**(6), 1–4 (2006)

7. Howe, J.: Crowdsourcing: How the Power of the Crowd is Driving the Future of Business. Random House, New York (2008)

8. Kanatani, N., Sasama, T., Kawamura, T., Sugahara, K.: Development of bus location system using smart phones. In: Proceedings of SICE Annual Conference, vol. 2010, pp. 2432–2433 (2010)

9. King, S.F., Brown, P.: Fix my street or else: Using the internet to voice local public service concerns. In: Proceedings of the 1st International Conference on Theory and Practice of Electronic Governance, pp. 72–80 (2007)

10. Schuurman, D., Baccarne, B., De Marez, L., Mechant, P.: Smart ideas for smart cities: investigating crowdsourcing for generating and selecting ideas for ict innovation in a city context. J. Theoret. Appl. Electron. Commer. Res. **7**(3), 49–62 (2012)

11. Shigihara, I., Arai, A., Saitou, O., Kuwahara, Y., Kamada M.: A dynamic bus guide based on real-time bus locations - a demonstration plan. In: 2013 16th International Conference on Network-Based Information Systems (NBIS), pp. 436–438 (2013)
12. Stembert, N., Mulder, I.J.: Love your city! an interactive platform empowering citizens to turn the public domain into a participatory domain. In: International Conference Using ICT, Social Media and Mobile Technologies to Foster Self-Organisation in Urban and Neighbourhood Governance (2013)
13. Watkins, K.E., Ferris, B., Borning, A., Rutherford, G.S., Layton, D.: Where is my bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. Transp. Res. Part A: Policy Pract. **45**(8), 839–848 (2011)