Evaluation of Pre-Acquisition Methods for Position Estimation System using Wireless LAN

Hiroshi Yoshida*, Seigo Ito** and Nobuo Kawaguchi***

*Graduate School of Information Science, Nagoya University
Furo-Cho, Chikusa-Ku, Nagoya, Aichi, 464-8601, Japan, yoshida@el.itc.nagoya-u.ac.jp
** Graduate School of Information Science, Nagoya University, seigo@el.itc.nagoya-u.ac.jp
***Graduate School of Engineering, Nagoya University

ABSTRACT

In recent years, position estimation systems using wireless LAN have been actively studied, along with the publication of many contributions. However, most of these location estimation methods require prior knowledge of wireless LAN access point locations. To deploy a location estimation system widely in the real world, it is crucial to develop an efficient pre-acquisition method of wireless LAN access point information. In this paper, we examine three kinds of pre-acquisition methods (gathering information by foot, bicycle, and car) and evaluate them according to the accuracy of the position estimation. Additionally, we present the acquisition performance of the efficient method we are now studying.

Keywords: Location Estimation System, Wireless LAN, RSSI.

1 INTRODUCTION

In establishing a ubiquitous computing environment, there are certain demands for positioning systems that can be used every time and everywhere. With the wide spreading implementation of wireless LAN, users can receive wireless LAN signals both indoors and outdoors. By using wireless LAN, many studies have been made on positioning systems [1][2][3][4].

Under these circumstances, we aim to construct a positioning system using wireless LAN as social infrastructure for everyone. To achieve this goal, the following three requirements must be met.

- **Everywhere**
  As a social infrastructure positioning system, it is necessary to have the ability to estimate positions everywhere. In other words, the system needs a wide-area positioning service.
- **Every time**
  Users want to know their position at any given time. Therefore, it is important to confirm the reliability and stability of a positioning system.
- **Every terminal**
  Our target devices for positioning include not only laptops, cell-phones, and PDAs but also embedded devices. These might be, for example, digital cameras with location information and budget trace tags.

This paper focuses on the first requirement.

The positioning system we assume uses a reference trace set consisting of *BSSID* (Basic Service Set Identifier), signal strength (*SS*), and position of access point (*latitude* and *longitude*). To construct a wide-area positioning system, it is important to collect the reference trace set effectively and widely.

Reference Trace Set = {*BSSID, SS, latitude, longitude*}

War Driving is one of the most famous methods of collecting a reference trace set by car. War Driving can collect reference trace sets widely and quickly. The Loki [5] project has collected 8 million reference trace sets and wigle.net [6] has collected about 7 million reference trace sets through War Driving. However, Kim's report [7] stated that there were some differences in reference trace sets between War Driving and War Walking. This difference can cause a difference in the abilities of wide-area positioning systems. It is thus necessary to examine the relationship between pre-acquisition methods and positioning accuracy. Little is known about pre-acquisition methods for wide-area positioning.

In this paper's work, we compared pre-acquisition methods for wide-area positioning from the viewpoints of the cost in transportation, the ease of acquisition, the time efficiency, the availability for diverse users, and the freedom of movement. Then, we conducted experiments on collecting reference trace sets by foot, bicycle, and car. By using these different reference trace sets, we evaluated positioning accuracy.

This paper is presented in four parts. First, we introduce related works and the Locky.jp project, which aims to construct a wide-area positioning system for Japan. Second, we compare pre-acquisition methods from various aspects. Then, we gather information in a real environment and conduct evaluation experiments. Finally, we present our conclusions.

.

2 RELATED WORKS

In this section, we introduce the methods, products and projects related to previously studied position estimation systems using wireless LAN.

2.1 Position Estimation Methods using Wireless LAN

Position estimation techniques can be divided roughly into Proximity [8], Triangulation [9] and Scene Analysis [10].

Proximity

Proximity is a technique that lists received signal strength from the access points observed by the estimation terminal and then estimates the position of the strongest signal access point as a position of the terminal. There is the disadvantage of lower position estimation ability than those of the other techniques, but it is possible to make estimates even in the past, since the estimation algorithm is easy and these are few access points.

Triangulation

Triangulation is a technique for estimating a position by using a relative position to the access point where the terminal observes the signal. The distance from the access point is calculated according to a wireless distance characteristic. Then, the position is estimated from the access point at three or more points.

Scene Analysis

Scene analysis is a technique that uses electrical wave reception strength to estimate a position by sampling in advance; this technique then uses the data acquired at two or more positions in the estimation area. The main advantage of this technique is accuracy at several meters. However, it has the disadvantage of taking much time for prior sampling in the area where the system is used.

2.2 Position Estimation Systems

Several systems have already been commercialized. The Ekahau Positioning System [11] is based on scene analysis and requires a position estimation model of the desired area by using a map that relates to the signal strength of indoor access points. AirLocation [12] estimates a position by using the difference in the arrival times of signals, and it requires recording of the area's reception signal strength prior to use. This system also requires a specific access point for measuring the time. SmartLocator [13] utilizes an infra-red tag with a wireless LAN to obtain fine-grained position estimation. Most of the current commercialized systems are for indoors and not for wide areas or outdoors. If we have a huge common reference trace set, these technologies might also be applied to wide area systems outdoors.

2.3 Acquisition Methods for Position Estimation System using Wireless LAN

In general, a reference trace set is collected by War Driving. War Driving is an acquisition method for a position estimation system using wireless LAN, and it is suitable for collecting reference trace sets over a wide range. However, not all access points can be discovered by this method, nor can all reference trace sets be collected. The cause of this problem is thought to be the metallic frame of the car and the car's speed in movement.

Some research has evaluated War Driving for its accuracy in position estimation as a system constructed by collected reference trace sets [7]. This research has reported that the accuracy of a position estimation system by both War Driving and War Walking is higher than the accuracy of a position estimation system by War Driving only.

Furthermore, there has been research on an acquisition method using a direction antenna [14]. This method can obtain high-accuracy position information, but it needs more time to collect the reference trace set.

2.4 Wireless LAN Location System Projects

There are several projects that have aimed to research and develop a position estimation system using wireless LAN [15][16].

PlaceLab [17] is a project to support development of a location information system by providing a toolkit for position estimation and a database for wireless beacon (IEEE802.11, GSM, and Bluetooth) information. PlaceLab's toolkit also supports collection of the reference trace set. PlaceLab has played a great role in the academic study of position estimation systems. However, currently, this project does not provide application level software for end-users.

Locky.jp [18] is our project designed to collect the open and common reference trace sets covering major locates in Japan through user collaboration. Locky.jp also aims to provide end-user location information services and applications. However, in the startup phase, it is difficult to expect ordinary users to collect reference trace sets by themselves without any reward. Therefore, in the initial phase, we decided to collect a certain number of reference trace sets by ourselves. After the initial collection, users who like our service and application may collect more data for us, so the initial data can enable a bootstrap process. Therefore, we need a cost- and time-efficient acquisition method of reference trace sets.

3 PRE-ACQUISITION METHODS

In general, a position estimation system based on wireless LAN estimates a position by using the database of the reference trace set. To deploy the system widely, we have to pre-acquire the reference trace set for a wide area. There is a variety of pre-acquisition methods based on transportation. As we explained in the previous section, the cost- and time-efficiency of the method is very important for our project. Additionally, the accuracy of the position estimation differs by the pre-acquisition method because the collected reference trace set differs by method. Accordingly, accuracy is also an important factor.

In the following, we explain the different pre-acquisition methods and their characteristics.

3.1 Pre-Acquisition Methods

In the following, we show the pros and cons of each pre-acquisition method performed by a mode of transportation. Table 1 summarizes this discussion.

- **Foot**

Walking is the simplest method of pre-acquisition. One can move anyplace by walking at an average speed of about 5 km/h. Therefore, one can acquire a reference trace set that is close to buildings and houses. This makes the signal stronger. However, the time efficiency of walking is the worst, so the overall score of walking is not good. In general, we hold the acquisition terminal with both hands in front of the body or with one hand at the side of the body.

- **Bicycle**

Generally speaking, the bicycle is a basic transportation method that can be easily used by everybody. Movement with a bicycle is naturally faster than on foot. The average speed is about 15 km/h. Therefore, the time efficiency of a bicycle is better than that of walking. Moreover, is also possible to collect information near buildings and houses. However, there is physical difficulty in riding a bicycle in hilly country. In general, the acquisition terminal is set in the front basket of a bicycle or placed in the user's backpack.

- **Motorcycle**

The motorcycle can move at almost the same speed of cars. It can easily perform wide-area acquisition. Compared with a car, a motorbike has the advantage of not getting jammed by traffic congestion, and it can easily enter narrow streets. During acquisition, movement speed is within 30 km/h. The acquisition terminal is placed in the user's backpack.

- **Car**

The pre-acquisition method using a car has the advantage of a very wide range of acquisition, as with a motorcycle. Furthermore, there is little physical load on the user, so it can easily perform a long-term acquisition. The main disadvantage of acquisition by car is that it can be influenced by traffic jams and regulations (e.g. one-way roads). Moreover, it is necessary to move at appropriate speeds according to the surrounding cars, so it cannot move freely to acquire ideal data. During acquisition, average speed is around 30 km/h or more according to the traffic situation. The acquisition terminal is set on the car's dashboard and fixed by tape.

3.2 Other Considerations

Received signal strength can be influenced by many objects such as a human body near the device. Thus, the transportation mode is not the only factor that affects signal strength. It has been reported that the intensity of the signal strength that can be observed at the same point is not stable. Furthermore, the state of the antenna at the time of acquisition and the surrounding environment also influence on the signal strength [9].

The date/time and the weather are also external factors affecting signal strength. For example, in daytime, there are many factors that may interfere with signal strength, such as people and cars, on the other hand, these factors decrease at midnight. In a residential area or a business district, the influence on the signal strength may differ depending on whether data is taken on a weekday or holiday. Moreover, atmospheric pressure, temperature, and humidity are related to a signal's propagation, and it is thought that these factors also influenced signal strength.

The position on the road may also affect the signal strength. For example, when collecting information by car, one cannot acquire data on a pedestrian sidewalk.

4 EXPERIMENT ON PRE-ACQUISITION

We conducted an experiment on the efficiency of the pre-acquisition methods and the influences of the mode of transportation.

4.1 Configuration of the Experiment

We collected reference trace sets by using three kinds of transportation (foot, bicycle, and car). Figure 1 shows the variety of our methods. We commonly use the wireless LAN card "Buffalo WLI-CB-AG54" and GPS "Garmin Geko 201" for the acquisition terminal. The left picture shows the style of the pre-acquisition by walking (on foot). The user keeps the acquisition device in front of him at chest height. The center shows the configuration of the bicycle acquisition. We placed the acquisition device in front of the handlebars. The right shows the configuration of the car acquisition. We placed the acquisition device in the passenger seat. The GPS unit is placed on the dashboard for better reception of the GPS signal.

Table 1: Qualitative Characteristics of Pre-Acquisition methods

	Foot	Bicycle	Motorcycle	Car
Cost of transportation	○	○	△	×
Ease of acquisition (physical load)	×	△	○	○
Time efficiency (range / time)	×	△	○	○
Availability (who can use it)	○	○	×	△
Freedom of movement (fine-grained)	○	△	×	×

Figure 1: Variety of Pre-Acquisition methods

To evaluate the accuracy and efficiency of each method under equal conditions, we restricted the configuration of the acquisition. We set the acquisition course as the outer circuit road of Nagoya University (2.0 km). Figure 2 shows this acquisition course [19]. The arrow in the figure shows the starting point and the direction. The course contains a variety of environmental features such as university buildings, a high-school building, a pond, and a residential area.

First, the reference trace set is collected as learning data by using each acquisition method. The data acquisition is performed every second, and the data of 1800 points (30 minutes) are collected for each method. As we restrict the acquisition time, we can round the course several times. For example, we round the course 3.5 times by bicycle. Next, another 1800 points for each method are also collected independently from the learning data as an experimental data for the evaluation. We evaluate the accuracy of the each pre-acquisition method using the GPS data of the experimental data as the reference position. We define “Accuracy” as the average distance between the estimated position and the reference position.

4.2 Access Point Location Acquisition

We performed data acquisition for two days. The basic results of acquisition are shown in Table 2. There is a difference in the acquired unique access point number between the learning data and the experimental data. The reason for the difference is the effect of the considerations described in Section 3.2. However, we use larger data as a learning data for each method. For both type of data, the bicycle data is the best in coverage rate of AP (over 90%). We define “Mean Time to AP Discovery” by the round time divided by AP count. This value shows that the AP discovery efficiency of the method using a bicycle is the best. Figure 3 shows the Venn diagram of the APs in the learning data. The set acquired by bicycle is the largest, and the set by car is nearly equivalent to the other methods. It takes 20 minutes on foot to round the course but less than 10 minutes by bicycle. The acquisition by car is not so fast due to the traffic signals. As we restrict the course, the acquisition is performed for multi-rounds. Figure 4 shows the cumulative number of APs through the acquisition. We can identify each round of each method from the figure. The result shows that it is always difficult to cover all APs by using a car. Also, the method using a bicycle is shown to be better than other methods for collecting APs. A part of the actually collected reference trace set log is shown in Figure 5. We adopt a compatible format with PlaceLab to accommodate future collaboration.

Table 2: Basic Results of Acquisition Experiment

		Foot	Bicycle	Car	Total
Learning data	Number of acquired unique APs	247	**269**	183	295
	Coverage rate of APs	83.7%	**91.2%**	62.0%	100%
	Time required for first round (sec)	1125	591	**486**	
	APs acquired by first round (rate of acquisition)	**214 (86.6%)**	209 (77.7%)	122 (66.7%)	
	Mean Time to AP Discovery (sec)	5.3	**2.8**	4.0	
	Average speed (km/h)	6.4	12.2	**14.8**	
Experimental data	Number of acquired unique APs	185	**256**	131	271
	Coverage rate of APs	68.3%	**94.5%**	48.3%	100%

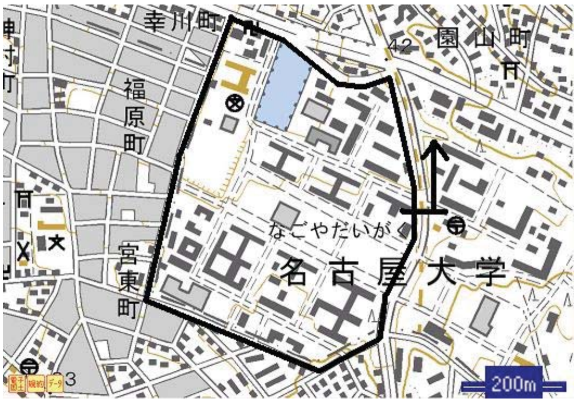

Figure 2: Outer Circuit Road of Nagoya University

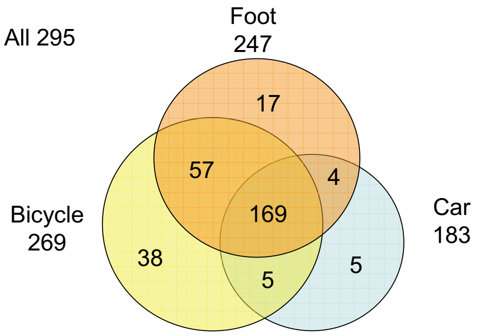

Figure 3: Venn diagram of APs in Learning Data

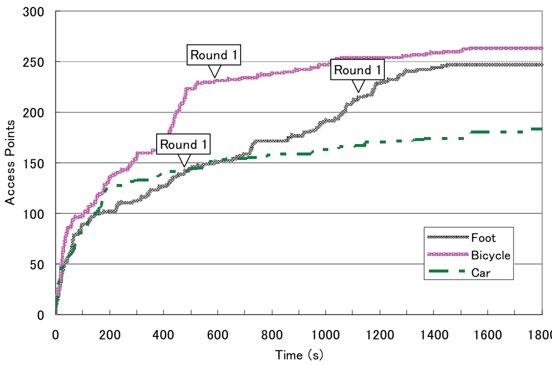

Figure 4: Cumulative number of collected APs


```
TYPE=GPS
  |TIME=1128332053640|HDOP=4.0|DGPSAGE=|DATE=031005
  |QUALITY=1|LAT=35.15483999999999|DGPSID=|STATUS=A
  |NUMSAT=00|MODE=A|ANTHEIGHT=44.8|SOG=1.9|VARDIR=W
  |VAR=6.9|LON=136.965385|TIMEOFFIX=093418
  |GEOHEIGHT=35.7|COG=140.4
TYPE=WIFI
  |TIME=1128332055625|ID=00:07:40:**:**:**
  |NAME=nuwnet|RSSI=-91|WEP=false|INFR=true
TYPE=WIFI
  |TIME=1128332055625|ID=00:07:40:**:**:**
  |NAME=nuwnet|RSSI=-90|WEP=false|INFR=true
```


Figure 5: Sample of Reference Trace Set

4.3 Analysis and Evaluation

We analyze the acquired data by estimating their positions using the experimental data. Table 3 shows the average position estimation accuracy (distance between reference data) and the coverage (rate of points that can be estimated). While using all data, we obtain the overall efficiency of each method. From this table, "foot" is the best for accuracy, but "bicycle" is the best for coverage. However, usually we don't acquire the data of the same place several times. Therefore, we analyzed accuracy by using data of the first round. The second group in the table shows the results. Coverage results are less than those for all data because of the smaller number of learned APs. However, accuracy is better than using all data, and the best is 23.2 m by using the bicycle data. This means that the smaller number of APs is better than full data. We think this is because of the signal strength. Thus, we analyzed the strong signal data (RSSI > -90 dbm) of the first round. The third group in the table shows the results. Coverage results also become worse, but this is not so bad in the bicycle data. The number of APs and the accuracy of the bicycle is also the best. This means that the method using the bicycle can find good signals efficiently.

Figure 6 shows the cumulative results from the experimental data for the method by foot. The accuracy is shown in the horizontal axis and the cumulative rate is shown in vertical axis. As this figure shows, the position estimation using the learning data acquired by foot can make estimates at less than 30 m in 68% of the area. The estimation data by car is worse by about 5%-10% in all ranges. Comparing the data by bicycle and by foot, there is about a 5% difference in the range of 10 m, but beyond 40 m, the bicycle data can estimate almost as well as that obtained by foot. Figures 7 and 8 show the results of using the experimental data based on the methods of bicycle and car. In each result, the cumulative rate of the bicycle data is better in the 40-m~ accuracy range.

Now we can select a pre-acquisition method. If we simply require accuracy, "foot" may be the answer. However, we need also coverage and efficiency in the acquisition. If we only think about speed, "car" may be the answer. But the accuracy and coverage of the car data are always the worst. Additionally, the cost of the transportation and lack of freedom are not acceptable for our project. As a result, it is clear that the method using a bicycle provides excellent efficiency for the pre-acquisition of the reference trace set.

Table 3: Position Estimation of Experimental data (by foot)

		Foot	Bicycle	Car
All data	No. of APs	247	**269**	183
	Accuracy (m)	**24.3**	26.7	29.6
	Coverage (%)	86.8	**87.8**	83.3
First-round data	No. of APs	**214**	209	122
	Accuracy (m)	31.3	**23.2**	26.0
	Coverage (%)	72.2	**74.6**	68.2
Strong signal APs (1^{st} round) (RSSI > -90)	No. of APs	155	**178**	49
	Accuracy (m)	30.1	**23.5**	32.9
	Coverage (%)	68.1	**73.2**	53.8

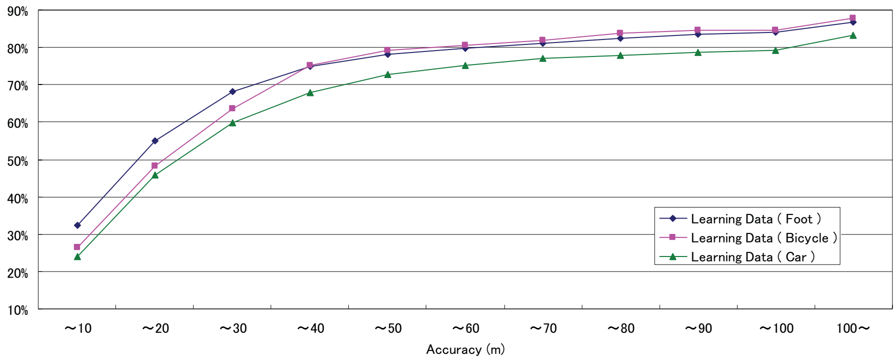

Figure 6: Cumulative Estimation Accuracy of Experimental data by Foot

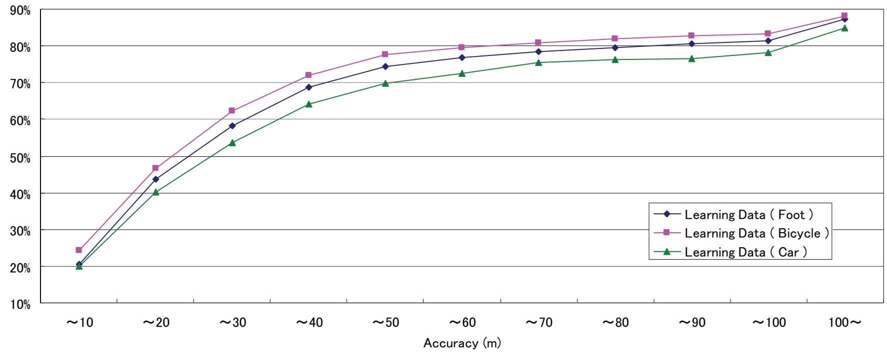

Figure 7: Cumulative Estimation Accuracy of Experimental data by Bicycle

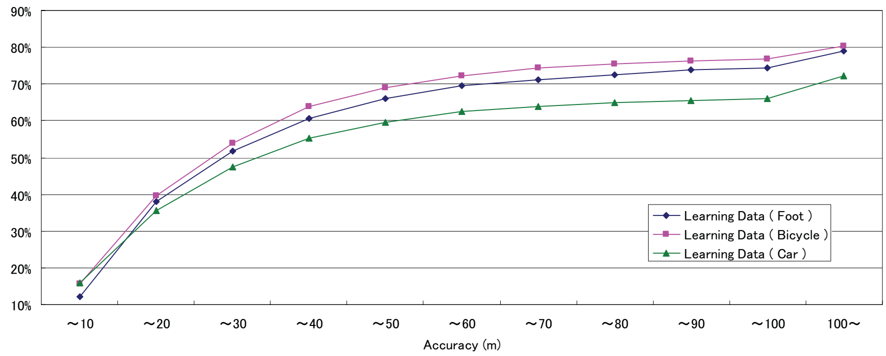

Figure 8: Cumulative Estimation Accuracy of Experimental data by Car

4.4 Differences in Estimation Methods

Finally, we confirm the dependency on position estimation methods. The results described above are based on the Proximity technique. We analyze the data using the Centroid technique, which is a technique of Triangulation. It estimates the position as the center of gravity of received APs. Figure 9 shows the results of the average accuracy of all combinations. The Y-axis is the accuracy and the x-axis is the acquisition method of the learning data. The six lines show combinations of the experimental data and the estimation method. From this figure, there is no big difference between Proximity and Centroid. Thus, there is no dependency on the technique used. Additionally, the method using a bicycle always has the best accuracy in the Centroid technique.

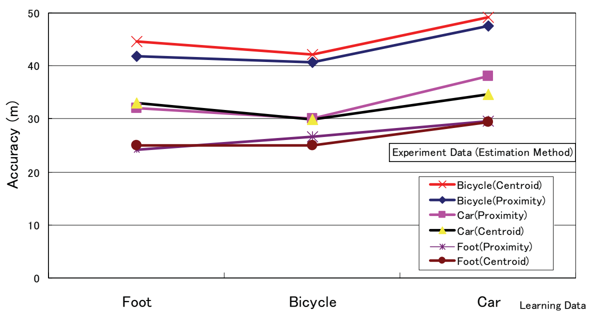

Figure 9: Difference between Proximity and Centroid

5 PRE-ACQUISITION TOOLS AND CURRENT STATUS

According to the results of the experiment, the pre-acquisition method using a bicycle shows excellent efficiency in reference trace set acquisition. We thus decided to use a bicycle for the wide-area pre-acquisition task in our project. To support real-world pre-acquisition using a bicycle, we have created the reference trace set acquisition equipment named the "Locky set," which is shown in Figure 10. The Locky set is composed of devices required for pre-acquisition, such as a notebook PC, GPS, and a wireless LAN card, assembled in a backpack. We also developed software named “Locky Stumbler,” which is designed for pre-acquisition of the reference trace set. Locky Stumbler has a sound output facility for monitoring the status of the acquisition. Additionally, Locky Stumbler can be controlled through an external controller that can be fixed to the bicycle’s handlebars. Thus, one can monitor and control the pre-acquisition while cycling without getting off from the bicycle. This system can continue the pre-acquisition for about 4 hours without changing the battery. Usually, we include an extra batte ry in the backpack, so the system can operate for 8 hours (nearly one workday).

Locky.jp is currently collecting a wide-range of reference trace sets using the Locky set. As an example of the pre-acquisition results, an example of the discovered access points is shown in Figure 11. This is a map of the 1-km-square area of the Sakae district in Nagoya, Japan. This area is in the city center of Nagoya surrounded by tall buildings. Each triangle icon indicates an acquired access point. In this area, we found 8.8 access points per 100 square meters.

Although still in the initial stage, we have already collected over 70,000 reference trace sets in Japan. We will continue our pre-acquisition activity to cover larger areas, especially in the major cities of Japan.

Figure 10: Locky set

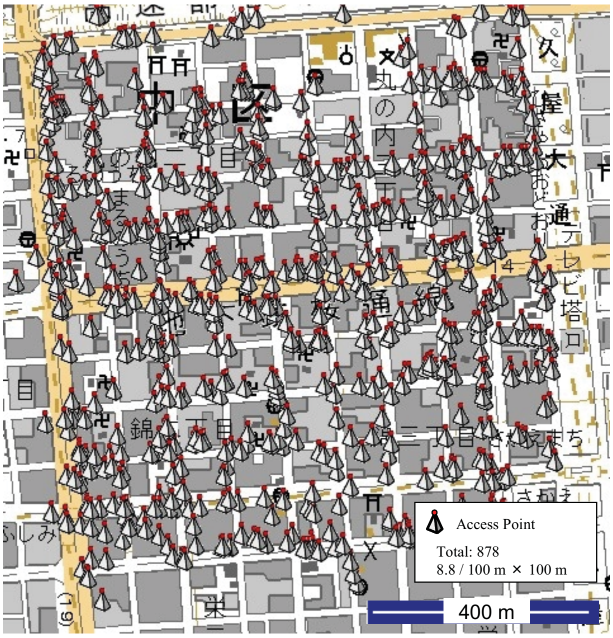

Figure 11: Example of collection results by Locky set

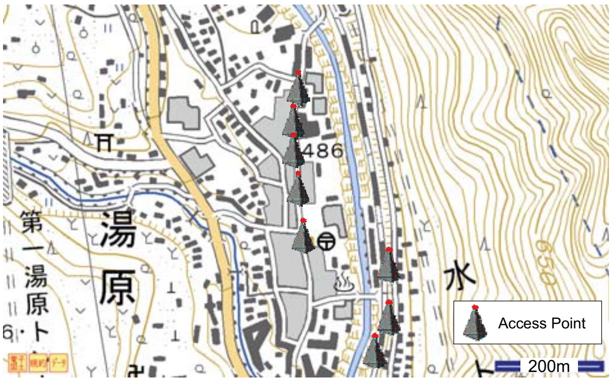

Figure 12: Example of collection result from rural area

Next, the results of collection in a rural area are shown in Figure 12. Here, the access points are separated by about 100 meters. The rural area has few access points, and it is difficult to use the Locky.jp system there. Therefore, it is necessary not only to collect reference trace sets but also to solve the problem of few access points. If the WiFi system becomes a major location system, an arrangement of access points used only for location accuracy will be deployed widely.

6 CONCLUSIONS

In this paper, we showed that the accuracy and efficiency of a pre-acquisition method are very important for the development and deployment of a wide-area position estimation system. We have performed detailed experiments to evaluate the accuracy and efficiency of pre-acquisition methods using three different modes of transportation. As a result, the pre-acquisition method using a bicycle is the best because it can collect more access points than other methods in the same time, and its accuracy is also high. To make pre-acquisition easier, we have created the "Locky set," which incorporates all required devices in a single backpack. By using Locky set, we are now collecting the reference trace sets for wide-area locations.

Further research is required on position estimation methods. Our Locky.jp database will be helpful for this work as a reference set. We also require an update method for the reference trace set. In our experience, access points in the field are nearly always stable, but some of them tend to change. Therefore, a reference trace set requires some kind of update periodically. One approach to developing an update method without GPS data would be to enable users to constantly update the position estimation system through collaboration.

REFERENCES

[1] Jeffrey Hightower, Sunny Consolvo, Anthony LaMarca, Ian Smith and Jeff Hughes, Learning and Recognizing the Places We Go, In Proceedings of the Seventh International Conference on Ubiquitous Computing (Ubicomp 2005), pp. 159-176 (2005)

[2] Paramvir Bahl and Venkata N. Padmanabhan, RADAR: An In-Building RF-based User Location and Tracking System, In Proceeding of The 19th Conference of the IEEE Communications Society (INFOCOM 2000), pp. 775-784 (2000)

[3] Anthony LaMarca, Yatin Chawathe, Sunny Consolvo, Jeffrey Hightower, Ian Smith, James Scott, Tim Sohn, James Howard, Jeff Hughes, Fred Potter, Jason Tabert, Pauline Powledge, Gaetano Borriello and Bill Schilit, Place Lab: Device Positioning Using Radio Beacons in the Wild, In Proceedings of The Third International Conference on Pervasive Computing (PERVASIVE 2005), pp. 116-133 (2005)

[4] Teruaki Kitasuka, Tsuneo Nakanishi and Akira Fukuda, Wireless LAN based Indoor Positioning System WiPS and Its Simulation, In Proceedings of 2003 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM 2003), pp. 272-275 (2003)

[5] Skyhook Wireless
- http://www.skyhookwireless.com/

[6] WiGLE
- http://www.wigle.net/

[7] Minkyong Kim, Jeffrey J. Fielding, and David Kotz, Risks of using AP locations discovered through war driving, In Proceedings of The 4th International Conference on Pervasive Computing (Pervasive 2006) pp. 67-82 (2006)

[8] John Krumm and Ken Hinckley, The NearMe Wireless Proximity Server, In Proceeding of the Sixth International Conference on Ubiquitous Computing (Ubicomp 2004), pp. 283-300 (2004)

[9] Anthony LaMarca, Jeffrey hightower, Ian Smith, and Sunny Consolvo, Self-Mapping in 802.11 Location Systems, In Proceedings of the seventh International Conference on Ubiquitous Computing (Ubicomp 2005), pp. 87-104 (2005)

[10] Jie Yin, Qiang Yang, and Lionel Ni, Adaptive Temporal Radio Maps for Indoor Location Estimation, In Proceedings of The Third IEEE International Conference on Pervasive Computing and Communications (PerCom 2005) pp. 85-94 (2005)

[11] Ekahau, Inc: Ekahau Positioning System
- http://www.ekahau.com/

[12] HITACHI AirLocation TM.
- http://www.hitachi.co.jp/

[13] NEC SmartLocator
- http://www.nec.co.jp/

[14] Hirokazu Satoh, Seigo Ito and Nobuo Kawaguchi, Position Estimation of Wireless Access Point using Directional Antenna, In Proceedings of The First International Workshop of Location and Context-Awareness (LoCA 2005), pp. 144-156 (2005)

[15] PlaceEngine
- http://www.placeengine.com/

[16] Kismet
- http://www.kismetwireless.net/

[17] PlaceLab
- http://placelab.org/

[18] Locky.jp
- http://locky.jp/

[19] Cyber Japan
- http://cyberjapan.jp/