Method for Introducing Wi-Fi-based Area Detection System

● Kazuo Hida ● Kensuke Sawada ● Yukio Hirokawa ● Yoshinori Yaginuma

The demand for mobility services that are available anywhere is increasing as work styles become increasingly diversified. Location-based services (LBSs) that leverage the user's present location to provide optimal services are attracting attention as a special class of mobility services, and to achieve indoor LBSs, a variety of techniques for detecting the area in which a user is present have been proposed. Area detection by wireless LAN (Wi-Fi) is particularly attractive since its use eliminates the need for installing a new infrastructure for positioning purposes. However, specialized knowledge has so far been required to judge the detection performance of a system using wireless technology. To eliminate this requirement, we have developed a method for visualizing inter-area detection performance so that non-specialists can judge detection performance of a Wi-Fi-based area detection system. This paper describes the method and shows the results of an evaluation test.

1. Introduction

The increasing diversification of work styles in recent years is driving a demand for mobility services that enable people to receive services just about anywhere. Such mobility services depend on the use of smart mobile devices and wireless access. Nowadays, the use of smartphones and wireless LAN (Wi-Fi) is spreading throughout the world,[1] and infrastructures supporting mobility services are being built.

Location-based services (LBSs), which provide services on the basis of the user's present location, are a class of mobility services that are attracting particular attention and expanding rapidly in the market.[2] Prominent LBSs include Foursquare, MyTown, Facebook Places, Loopt, and Gowalla.[3]

For LBSs in outdoor environments, GPS is often used to obtain location information while for indoor environments, a variety of technologies have been proposed, such as Wi-Fi, ultrasonic, infrared,[4] and the Indoor Messaging System (IMES) using GPS technology.[5] However, applying new technologies in an indoor environment requires the installation of a new infrastructure, the cost of which has so far hindered the introduction of indoor LBSs. As a consequence, attention has turned to the existing Wi-Fi infrastructure, which has found widespread use throughout the world. Using an existing infrastructure can reduce the number of hurdles that need to be overcome before introducing LBSs. In addition, using technologies commonly used worldwide can reduce the number of hurdles standing in the way of global LBS expansion.

One hurdle standing in the way of effective operation of indoor LBSs is multipath interference caused by walls and other obstacles. This is because it degrades the accuracy of Wi-Fi-based systems that detect the area in which a user is present. This led to the development of an area detection technology robust to complex propagation environments with multipath interference. This technology first obtains the radio features of different areas beforehand by sampling the received signal strength indicators (RSSIs) of the Wi-Fi access points (APs). Then, when determining the location of a particular user, it infers that the user is present in the area for which the registered features are most similar to the features presently being observed (**Figure 1**). This scheme is attracting attention as an area detection technology that can be deployed in a customer's environment and that is relatively unaffected by environmental changes. Yet, given the characteristics of Wi-Fi signals, specialized knowledge in radio propagation is needed to judge the area detection performance for the target environment, and this requirement has also hindered the introduction of LBSs.

To solve this problem, we have developed a method for introducing an indoor area detection system that enables non-specialists to evaluate the detection performance for an area through visualization of that performance. In this paper, we introduce the method and present the results of an evaluation test.

2. Conventional methods

One conventional method of judging the performance of area detection using Wi-Fi is to conduct an analysis based on visualizing the RSSI of a Wi-Fi AP. The Wi-Fi RSSI at different points is measured using a terminal, and the signal strength is displayed in grayscale on a map (**Figure 2**). This enables area detection performance to be analyzed through the visualization of radio propagation conditions.[6] However, analyzing area detection performance from this information still requires specialized knowledge of radio propagation.

There are also products that aim to maintain a certain level of detection performance by specifying, in the instruction manual, the infrastructure installation intervals, installation environments, etc. needed for effective operation. However, preparing instruction manuals tailored to all customer environments is difficult. Moreover, it is also difficult to judge detection performance simply by explaining installation intervals and installation environments. There is also a need to support APs that have already been installed.

Other related technologies involve the checking of detection performance by evaluation testing after the detection environment has been constructed. However, many tests are needed to gather a sufficient amount of evaluation data, which can delay the introduction of a LBS system. In short, there is a need for a method that requires no specialized knowledge of radio propagation and that enables non-specialists to easily judge area detection performance.

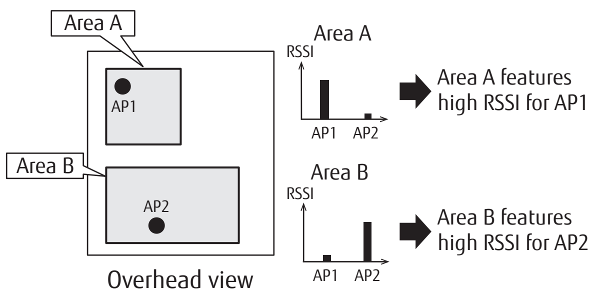


Figure 1
Area detection by Wi-Fi.

3. Proposed method

To enable non-specialists to easily judge area detection performance, we developed a method for making a logical evaluation of detection performance. It divides the performance-judging process into five steps.

1) Input environment information

A Wi-Fi-based area detection method can use a customer's environment effectively. It is therefore necessary to acquire information on the Wi-Fi signal environment at the target site and to evaluate whether the customer will be able to detect the areas desired. This step converts information on those areas and on existing APs into data that can be used to evaluate area detection performance in the customer's environ-

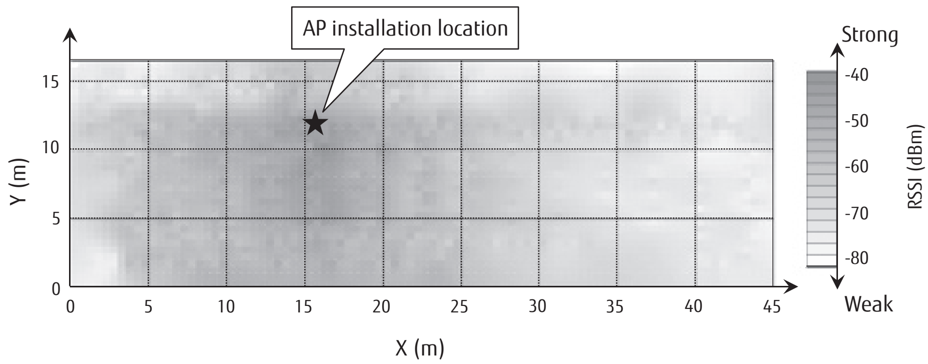


Figure 2
Visualization of radio propagation conditions.

ment. It targets, in particular, information related to areas and existing APs for data conversion.

- Definition of each area
- Floor space of each area
- Installation position of each AP
- Transmission power of each AP
- Transmission frequency of each AP

A graphical user interface (GUI) consisting of an environment-information input screen is used to enter the above information into the system (**Figure 3**). The user operates a mouse to drag and drop rectangles representing the definition and size of areas and solid circles representing existing APs onto a map of the customer's site.

2) Simulate area detection accuracy

A proprietary algorithm is then used to calculate and visualize the area detection accuracy for each user-defined area on the basis of the environment information entered in step 1).

The first task in this step is to calculate theoretical received signal strength distribution characteristics in all areas for each AP on the basis of the AP installation position, transmission power, and transmission frequency entered in step 1). The propagation-loss estimation method for indoor environments described in an ITU-R recommendation[7] is used to calculate the theoretical value. The next task is to obtain an RSSI vector at every point by using the calculated theoretical value. Then the similarity of RSSI vector is calculated by using vector correlation. Two areas for which the obtained RSSI vectors are dissimilar are easy to distinguish, and conversely, two areas with similar RSSI vectors are difficult to distinguish. We can therefore define the area detection accuracy at a certain point as the size of the area having RSSI vectors for which the similarity with the RSSI vector at that point is above a threshold value. The similarity of RSSI vectors is determined by vector correlation. Finally, area detection accuracy as determined by the above calculations is superposed on the area map entered in step 1) and then output.

An example of the output is shown in **Figure 4**. A point at which area detection accuracy is low is shown in dark gray while a point at which it is high is shown in light gray. As shown in Figure 4, several areas have many light gray-shaded points and thus can be expected to have high area detection accuracy. A few contain dark gray-shaded points and thus can be expected to have low detection accuracy.

3) Measure strength of received signal

The RSSI data needed to evaluate performance are collected by having personnel walk uniformly throughout each area for a specific period of time while carrying terminals that measure the RSSI. This measurement time for each area is calculated using the floor space entered in step 1). The formula below is used to calculate measurement time T (min) in terms of area floor space S (m^2), measurement-terminal sampling interval t (s), and number of terminals n. The coefficient 0.103 was determined by the result of preliminary experiment. This formula is used to calculate the amount of time needed for obtaining data on all possible positions and orientations that a person could assume within a particular area.

$$T = 0.103 \times S \times t \div n$$

4) Visualize Wi-Fi signal environment

The RSSI data collected in step 3) are used to determine the features of each area's Wi-Fi signal environment and to visualize the similarities of environments between those areas. This enables identification of those areas susceptible to misdetection and of those

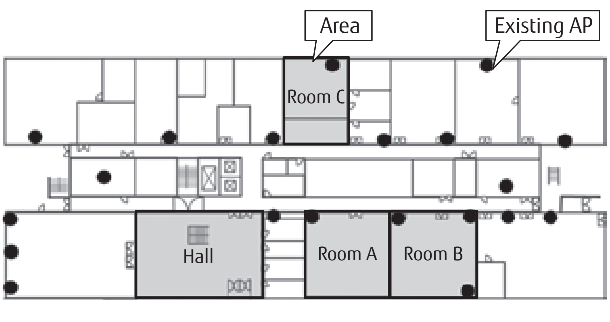


Figure 3
Environment-information input screen.

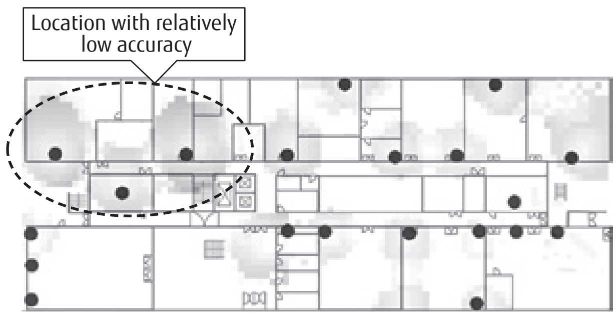


Figure 4
Simulation results.

areas for which area detection performance can be guaranteed.

The correlation coefficient (R) between the RSSI data (x_1, x_2, ..., x_n) for a certain area (X) and the RSSI data (y_1, y_2, ..., y_n) for another area (Y) is calculated using

$$R = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

where $\bar{x}$ and $\bar{y}$ are average RSSI in the area (X) and area (Y), respectively.

The R is calculated for each area pair, and a histogram of the R values is prepared. The shapes of the histogram can be classified into five patterns on the basis of the position and heights of the frequency peaks, as shown in **Figure 5**. If the frequency peaks for two areas have a low correlation coefficient, their Wi-Fi signal environments are considered sufficiently different. Therefore, area detection performance is expected to be high.

5) Calculate misdetection rate

Finally, area detection performance is visualized by calculating the area misdetection rate for each area using some of the RSSI data collected in step 3). The RSSI data are divided into training data and observed data, and area detection is executed. The number of misdetections for each area is calculated. The misdetections rate is the average of the probability that the target area is falsely recognized as another area and the probability that another area is falsely recognized as the target area. The misdetected areas are thereby identified.

This determination of area detection performance directly from the misdetection rate and logically from the similarity of Wi-Fi signal environments enables area detection performance to be judged from more than one viewpoint. Furthermore, since area detection performance in either case is determined from actual measurements, reliable area detection performance can be guaranteed.

4. Evaluation

We tested area detection performance using the proposed method assuming the layout shown in **Figure 6** for the customer's worksite. In this test, we used a total of nine APs installed on the same floor for area detection and defined four areas (meeting room A, meeting room B, office, and hall).

We used a proprietary pattern-matching algorithm[8] developed by Fujitsu Laboratories for area detection. This algorithm uses a probability distribution that considers the balance between hardware requirements/performance and area detection accuracy. We carried out signal measurement by walking through each area for five minutes using two smartphones, each equipped with an RSSI data collection app (sampling interval: 0.5 s). We calculated this signal measurement time using the formula presented in step 3) of the previous section.

The results of calculating area detection accuracy as described in step 2) of the previous section are shown in Figure 6. As shown, the office area includes many light gray-shaded points, which means that this area can be presumed to have good detection accuracy. On the other hand, meeting room A, meeting room B, and the hall each includes dark gray-shaded points, which means that these areas can be presumed to have poor detection accuracy.

The frequency distributions of the correlation coefficients calculated using the procedure described in step 4) of the previous section are shown in **Figure 7**. In distributions (a), (b), and (c), the frequency peak

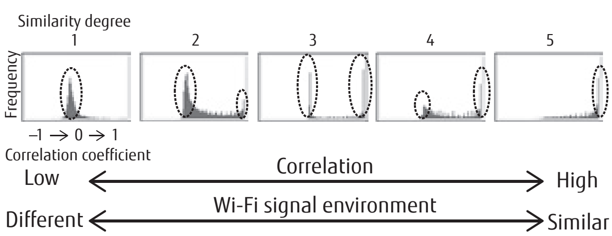


Figure 5
Patterns of correlation coefficient distribution and correlation of Wi-Fi signal environments.

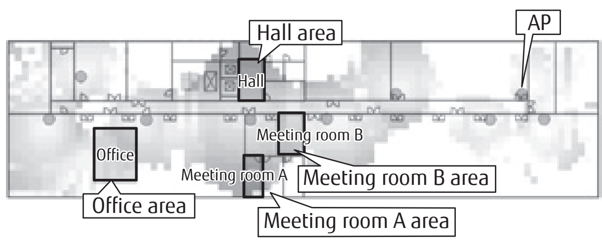


Figure 6
Evaluation test environment and results of simulating area detection accuracy.

lies very close to a correlation coefficient of 1 while the frequency peak lies near a correlation coefficient of 0 in (d), (e), and (f).

The frequency distributions can be matched with the patterns in Figure 5 to classify each area in terms of degree of similarity. Specifically, distributions (a), (b), and (c) resemble similarity degree 5 in Figure 5 while (d), (e), and (f) resemble similarity degree 1. These results are summarized in **Table 1**. As shown, the office area can be classified as similarity degree 1 with respect to all other areas, meaning that this area has a Wi-Fi signal environment different than that of the other areas. For this reason, it can be inferred that no misdetections will occur in the office area. In contrast, the meeting room and hall areas are each classified as similarity degree 5 with respect to each other, which means that the Wi-Fi signal environments of these three areas are similar. It can therefore be inferred that misdetections will easily occur in these three areas.

The results of determining the misdetection rate for each area pair using the procedure described in step 5) of the previous section are summarized in **Table 2**. The misdetection rates for the office area are 0%, which matches the results for degree of similarity in Table 1. However, while it can be inferred from the results of Table 1 that misdetections will occur among the meeting room and hall areas, Table 2 indicates that no misdetections will occur between the meeting room A and hall areas. The reason for this is thought to be that the measurements we performed collected data for which a misdetection is unlikely to occur. The maximum misdetections rate was about 3%, which is low overall.

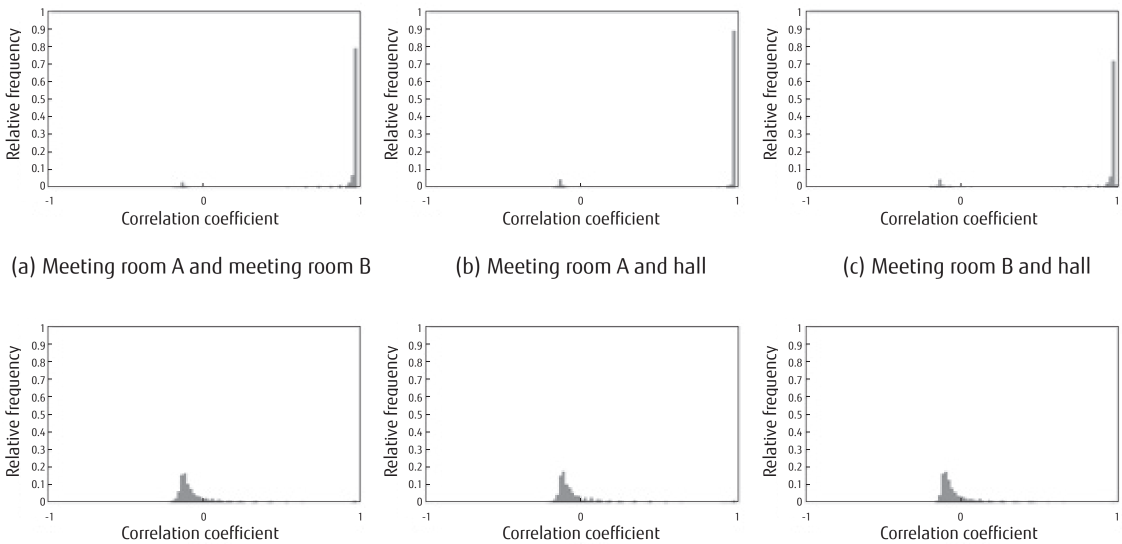
From the results given in Tables 1 and 2, the customer can be told that area detection will have high accuracy in the office area and that misdetections will

Table 1
Degree of similarity of Wi-Fi signal environments for area pairs.

	Meeting room A	Meeting room B	Hall	Office
Meeting room A	–	5	5	1
Meeting room B	–	–	5	1
Hall	–	–	–	1
Office	–	–	–	–

Table 2
Misdetection rate for area pairs.

	Meeting room A	Meeting room B	Hall	Office
Meeting room A	–	3.3%	0.0%	0.0%
Meeting room B	–	–	1.6%	0.0%
Hall	–	–	–	0.0%
Office	–	–	–	–



(a) Meeting room A and meeting room B
(b) Meeting room A and hall
(c) Meeting room B and hall
(d) Office and meeting room B
(e) Office and hall
(f) Office and meeting room A

Figure 7
Frequency distributions of correlation coefficient between areas.

likely occur in the meeting room and hall areas.

When some areas have a high likelihood of misdetections, the area detection rate can be improved by redefining the areas or by adding more APs. In this evaluation, we redefined the areas as follows.

The results of Table 1 indicate that the Wi-Fi signal environments for the meeting room and hall areas are similar and that they differ from that of the office area. We can therefore integrate these three areas into one and define two areas: one consisting of this integrated area and the existing office area. By integrating areas having similar Wi-Fi signal environments and defining them as a single area, we end up with only areas for which the Wi-Fi signal environments are different, thereby improving the area detection rate.

5. Conclusion

Our proposed method clarifies the performance of Wi-Fi-based area detection systems that detect the indoor area in which a user is present. It is performed by a logical evaluation that divides the performance-judging process into five steps. It requires no specialized knowledge of radio propagation or extensive testing. Evaluation testing showed that areas for which detection accuracy will likely be high as well as areas where misdetections will likely occur can be identified using this method.

This method will enable non-specialists to easily judge the performance of Wi-Fi-based area detection—a technique that can be easily expanded throughout the world—while acting as a method that can reduce the number of hurdles that need to be overcome before introducing LBSs. We can expect the application of this method to stimulate the market for LBS mobility services.

Additionally, as the proposed method has a function for visualizing the features of a high-dimension vector in a specific space, it could also be applied to high-dimension vectors other than ones representing the characteristics of radio propagation. For example, if a vector containing elements representing various attributes of communication or work in an office was to be defined, this function could be used to analyze that office space and visualize practices that contribute to inefficient communication or work procedures. In short, it could be used to create a system for improving office layout and work efficiency.

References

1) Cisco: Cisco Visual Networking Index (VNI) Global Mobile Data Traffic Forecast, 2013–2018, 2014.
2) Andre Malm: Mobile Location-Based Services. Berg Insight, 2013.
3) S. Salt: Social Location Marketing: Outshining Your Competitors on Foursquare, Gowalla, Yelp & Other Location Sharing Sites, Pearson Education, 2011.
4) J. Hightower et al.: Location System for Ubiquitous Computing, IEEE Computer Society, Vol. 34, No. 8, pp. 57–66, 2001.
5) A. Dempster: QZSS's Indoor Messaging System, GNSS Friend or Foe, InsideGNSS, technical article, January/February 2009.
6) Ekahau: Ekahau Site Survey. *http://www.ekahau.com/wifidesign/ekahau-site-survey*
7) ITU-R: Propagation data and prediction models for the planning of indoor radio communication systems and radio local area networks in the frequency range 900 MHz to 100 GHz, ITU-R Recommendations, pp. 1238–7, 2012.
8) K. Hida et al.: Evaluation of Area Detection Method Using Machine Learning. Multimedia, Distributed, Cooperative, and Mobile Symposium 2014 (DICOMO2014), 8G-2, 2014.

Kazuo Hida
Fujitsu Laboratories Ltd.
Mr. Hida is engaged in the research of operation environments for ubiquitous systems targeting people in the real world.

Yukio Hirokawa
Fujitsu Ltd.
Mr. Hirokawa is engaged in the planning and development of the SPATIOWL location data service.

Kensuke Sawada
Fujitsu Laboratories Ltd.
Dr. Sawada is engaged in the research of wireless ubiquitous systems.

Yoshinori Yaginuma
Fujitsu Laboratories Ltd.
Mr. Yaginuma is engaged in the development of healthcare solutions and their elemental technologies.