A Proposal of Landmark-conscious Voice Navigation

Shota Watanabe†, Katsuhiko Kaji‡and Nobuo Kawaguchi*

†School of Engineering,Nagoya University
Fro-cho, Chikusa-ku, Nagoya, 464-8601, Japan, shon@ucl.nuee.nagoya-u.ac.jp
‡Graduate School of Engineering, Nagoya University, kaji@nuee.nagoya-u.ac.jp
* Graduate School of Engineering, Nagoya University, kawaguti@nagoya-u.jp

ABSTRACT

Recently, pedestrian navigation systems which display a guide route on the map become popular. However, the user needs to look at the device's display many times when using these navigation systems. It is necessary for user to stop for look at the map on the display, in order to walk safely,. Moreover, if an user looks at the display while walking, he is at risk for colliding something, or losing sight of the route. Therefore, we propose a landmark-conscious voice navigation. Landmark means something easy to be recognized, such as a billboard, a monument, or a vending machine. Our voice navigation method provides a guide message using highly visible landmark, based on information of landmark. The visibility is calculated by position, size, visible direction, and features of the landmark. To exemplify the feasibility of the voice navigation, we have implemented a prototype system.

Keywords: Pedestrian, Voice navigation, Landmark, Visibility

1 INTRODUCTION

Due to the proliferation of mobile devices equipped with GPS and geomagnetic sensors, navigation systems for pedestrians (such as NAVITIME [1], Google Maps Navigation [2], etc) are becoming widespread. Many of these navigation systems display maps, routes, and informational text on the terminal display and the user walks following the route and informational text. In these systems, the user needs to look at the display many times and verify that the route they are presently walking is correct. For safe walking, it is necessary to stop to look at the terminal. Also, when the user looks at the terminal while they are walking there is the danger of collision or the possibility of losing sight of moving landmarks. If it were possible to be guided to the destination solely by voice, it would be possible to look at the surroundings while listening to the guide. In directions premised on looking at a map, for the information that the map lacks, the user himself has supplement the information by searching out objects that can serve as landmarks from the map. However, with voice-only directions, it is assumed that the system will have to indicate the landmarks.

We present Landmark-conscious Voice Navigation. A landmark is something like a billboard, monument, or vending machine which serves as a visual guide for the pedestrian's movement. Since it is assumed in guides using landmarks that the pedestrian will use their movement position as the base, it is effective for grasping the present location or action points. In our voice-navigation method, using landmark data registered beforehand, it selects high-visibility landmarks at the pedestrian's location. The visibility of landmarks is calculated using the position of walls, the position of landmarks, their size, the directions from which they are visible, and other such characteristics. The informational text generated uses information such as the direction from which the pedestrian can see the landmark, its distance, text written on the landmark, and its color. Also, to verify the practicality of the voice-navigation we have proposed, we have implemented a prototype system.

In Part 2 below, we will discuss the research related to this study, in Part 3 we will propose Landmark-conscious Voice Navigation, and in Part 4, we will discuss the implemented prototype. Also, in Part 5, we will discuss the collection of landmark data to be used in navigation, and we will provide a summary in Part 6.

2 Related Research

In pedestrian navigation, which is premised on showing a map, a route search from the departure point to the destination is performed, and direction text and an deformed map are generated. [3] In generally-available maps, since there is too much information that can be acquired at once, the user himself must selectively acquire the necessary information. In an deformed map, since only information that is deemed necessary for the user to walk along the route, route errors can be reduced. When both an deformed map and information text is used simultaneously, the user follows the information text and walks, and if there is insufficient information in the informational text the user himself supplements the insufficient information from the deformed map. With this type of deformed map, the informational text provided in a navigational system premised on supplementing information, there is little information needed for selecting routes included in the informational texts in advance, and there are many simple directions such as " Turn left XX meters ahead."

Using the positional relationship between objects serving as landmarks, Fujii, et. al. [4] conducted research into ways to provide clearer guidance to pedestrian's actions. Fujii, et. al., by applying a structuralized model in which the structuralized objects are expanded on all the figures on the map, produce informational text including spatial expressions indicating the relative positional relationship with the objective. However, since outdoor space is the object, objects treated as landmarks are limited to the buildings to the side of the route or intersections. In outdoor space, the locations of objects like buildings are not limited to the side of the route or intersections. Also, since large buildings are used as landmarks, only the name of the landmark and its position are included in the informational texts, but landmarks necessary for supplemental information, such as landmarks with a distinctive color, may be useful for the informational text.

As navigation support which indicates the direction to proceed along optional points, there is the study by Tsukada, et. al.[5] Mounting directional sensors, GPS, and multiple actuators on a belt, they made tactile information display with directional information possible. It is superior insofar as it eliminates the hassle caused by navigation systems, but since it does not direct the user's next action in advance, it is difficult to provide guidance on what action to take next. In cases where the action to be taken next is not sufficiently relayed, the pedestrian ends up walking the way the think they should, resulting in route errors. For that reason, in voice-only navigation as well, it is necessary to provide guidance on the direction in which to proceed after the action has been taken in order to reduce route errors.

In studies such as this one, which consider the visibility of landmarks, Nakazawa, et. al., [6] paid attention to the visibility of landmarks. They produced an evaluation model in order to quantify the data, conducted route searches focusing on the quantified high-visibility landmarks, and created an informational map. In this study, since we focus our attention not on the creation of informational maps, but on voice-guides which consider the visibility of landmarks at the pedestrian's present location, we select high-visibility landmarks in optional locations along the designated route. For that reason, it is necessary to select high-visibility landmarks at the respective locations where guidance will be provided, not to select high-visibility landmarks along the entire route. Also, when calculating the visibility of landmarks at optional points, it is necessary to consider from which direction the landmark can be easily recognized. Therefore, in this study, we consider the possible directions from which the landmark is visible, and select highly-visible landmarks.

3 Landmark-conscious Voice Navigation

In this chapter, using structured map data, we will discuss a method which makes voice-only directions between the user-inputted point of departure and destination possible. Taking into consideration the points in Part 2, we will list the conditions for sufficiently understanding voice-only instructions on the premise that a map will not be looked at, which is the goal for this study.

- Give directions treating all high-visibility objects existing in the vicinity of the route as a landmark.
- Include directions regarding supplemental information for the identification of landmarks.

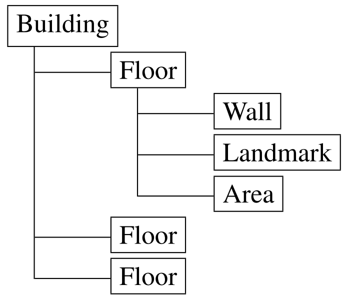

Figure 1: Structure of Building Data

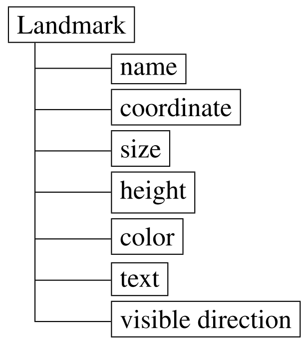

Figure 2: Structure of Landmark Data

- Give directions on the route to proceed in after action has been taken in order to reduce route errors after action has been taken.
- Give directions using high-visibility landmarks using optional points.

3.1 Structure of the Map Data

In order to create informational text, landmark information which is visible to the pedestrian is necessary. In order to extract landmarks that can be visible, map information and information regarding the landmarks are necessary. Below, we will discuss the necessary items for the information. When considering indoor space as the object, buildings form the basis of the map data. Data with buildings as the unit are presented in Figure 1.

Buildings consist of floors. Floor information consists of walls, landmarks and areas. Also, the route used in navigation lead with generally-used route estimations. The shortest route from the departure point to the destination is presented as the route to follow, and node and link information is used to derive the route.

The data structure of landmarks used in buildings is presented in Figure 2.

In order to calculate the visibility of landmarks from optional locations, the directions from which the landmark is visible is held as information. It is necessary to have supplemental information to identify the color, size, written text and height from the floor. In addition, region, area name and categorized area information are treated as landmarks. This is because things such as stairwells, rooms, and ticket gates can also become landmarks within the identified area in the region which can be walked in. This landmark information is comprehensively registered in the space which is the object of navigation. If the aforementioned information is available, it is possible to extract the landmarks visible from optional locations and their direction. Also, it is possible to calculate the possible visible distance using the size of the landmarks, their height from the floor and other such features.

3.2 Method of Selecting Landmarks at Optional Locations Used in Directions

Landmarks are used in giving directions in order to indicate the next location to aim for from a certain place. Therefore, landmarks should be visible from the point where the directions are given as well as exist near the place where the next action is to be taken. Also, if multiple landmarks be extracted, the selections should be landmarks that are easy to for the pedestrian to perceive. In order for the pedestrian to recognize the landmarks in the informational text, it is necessary to select landmarks which can easily enter the pedestrian's field of vision. "Exists in the direction the pedestrian is headed, and can be confirmed from a distance" and "is of a distinctive color," among others, can be given as features of landmarks which easily enter the pedestrian's field of vision. Based on the above, the selection of landmarks will be performed with the following three steps. The selection of landmarks shall be performed after the pedestrian's route is determined. Landmarks will be selected using Action Node P, which is one action before Action Node Q, which is along the route.

(A)Selecting Landmarks Visible from Node P
Using the present floor landmark list L and the wall list W, a list of landmarks which can be seen from Node P, L_1, will be created. From among the items in L_1, landmarks which exist in the direction from which Node Q progresses from Node P, and Node P exists within the range of angles visible shall be selected, and landmark list L_2 will be generated.

(B)Selection of Landmarks Existing in Positions Visible from Node Q
Using L_2, the landmark list selected in (A), and wall list W, a list of landmarks which can be seen from Node Q, L_3, will be generated.

(C)Selection of Most Highly Visible Landmark
From list L_3, the list of landmarks selected in (B), high-visibility landmarks L_Q are selected. The degree of visibility will be calculated from the proximity to Node P and the size of the landmark. Also, depending on the user in question, the color, shape, text, and height from the floor may affect the degree of visibility.

Landmark L_Q, which was obtained from the three steps above shall be the landmark selected in Node Q. Also, if multiple easily-recognized landmarks are extracted, if information is given as to their positional relationship, this will increase the effectiveness of the directions.

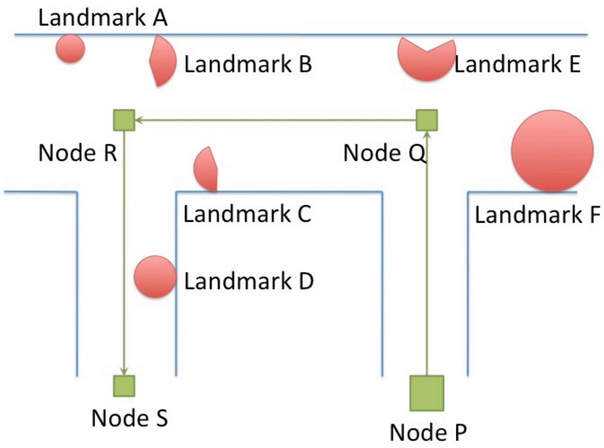

Figure 3: Landmark selection example

Examples of landmark selections are illustrated in Figure 3. The pedestrian is positioned in Node P, and when led in order to Nodes P, Q, R and S, the landmarks used to indicate the direction from Node Q to Node R will be selected from Landmarks A, B, C, D, E and F. The landmark data shall be the angles from which it is visible and its size, and will show the angles from which it is visible and its size from a fan-shaped angle and its size. Also, floors will be indicated by a straight line. First, by using the floor and landmark position information, Landmarks A, B, C, E and F, which are located in a position visible from Node Q, will be selected. Next, from these, Landmarks A and B, which exist in the direction of progression from Node Q to Node R, and exist in exists within the visible range of angles for Node Q, will be selected. From these, Landmarks A and B, which exist near R, will be selected. Finally, from this, Landmark B, which is the largest, will be selected. From this, providing direction with Node Q using B is possible, and with P, the directions will be generated in the order of "Turn left at the end, and head for B, which will be visible on the right when you advance a little forward" will be given, and afterward, "Turn left at corner B."

3.3 Generating Informational Text

In order to provide instructions of on the direction to head in after the action ahs been taken, in an informational text in a certain point, two directions will be performed. The point, the point of the next action, and the point of the action after that shall be expressed as P, Q, and R. The "point of the action" is the point where the pedestrian turns or the point (area) that is passed through. In the informational text at Node P, it shall be deemed as the next first half and second half.

- Directions on the actions from Node P to Node Q
- Directions of the direction to head toward from Node Q to Node R

In order to provide these directions, two landmarks selected in the respective points between two action nodes — the next action node and the action node following it — are used in order to generate informational text. Also, when providing directions for actions at a point where the pedestrian passes through an area, as the area itself fulfills the function of a landmark, it is used for information on the area, not as a landmark.

This method of generating informational text is described in Figure 4. Nodes P, Q and R in the diagram correspond to the present node, the next action node, and the action node after that. Landmarks A and B are the landmarks are those selected by Nodes Q and R respectively.

First, we will discuss the generation of informational text for the first half. Using progression direction component l_1, the distance between Node P and Landmark A, and the size of Landmark A, m_1, the distance that Landmark A can be seen from Node P, l_a, is calculated. s_a is calculated to see whether Landmark A is to the left or the right side of the segmented line joining Node P and Node R. Next, the movement direction component, l_2, the distance between Landmark A and Node R, is calculated. Finally, the positional relationship, p_a, between Landmark A and Node R seen from Node P is calculated. Directions are generated using this information and the features of Landmark A f_a.

In the generated informational text, l_a is used to indicate the distance up to the point the landmark can be seen, and s_a is used to indicate the direction in which the landmark can be seen. l_2 and p_a are used to indicate the directional relationship between the action points and the landmarks. Also, f_a is used to relay the landmarks in order to indicate the landmarks clearly. With this, we will now present an example of the directions generated. "When you walk a little, turn left at the corner beyond the red billboard visible to the right." When passing through an area such as a staircase straight ahead, the following informational text can be produced: " After going down the stairs, turn right in front of the elevator right directly in your path." Next, we will discuss the second half of the creation of informational text. Here, from the following action points, it is enough to be able to understand the direction to head in. Therefore, we will discuss from what distance, in what direction, and in what way Landmark B can be seen. Similar to the first half, we will find l_b,s_b and f_b and generate informational text in the same way as in the first half.

With this, the following informational text can be generated: "After that, proceed straight for a while in the direction of the green poster which says " Coffee," which can be seen on the right." In this way, in the first half, using the landmarks that were used one direction prior, directions for the next action are given, and in the second half, the direction to proceed in until the next action is indicated. With this, since the place to proceed to can be continuously grasped, the pedestrian can walk smoothly.

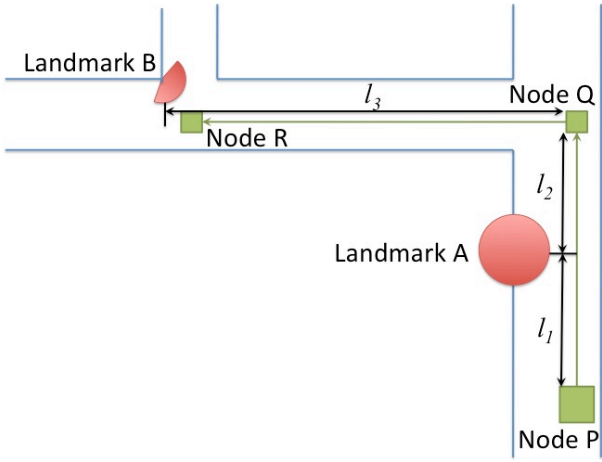

Figure 4: Directional Relationship when the Informational Text is Generated

4 Prototype Implementation

We implemented the navigation prototype examined in Part 3. In this system, information text which gives directions from the user-inputted point of departure and destination is created from floor and landmark information registered in advance. Also, in this system, there is no collaboration with position-gauging technology. Here, with the inside of the Nagoya University Station subway terminal as an example, we shall analyze the informational text generated by the prototype system.

Examples of informational text generated by this system are illustrated in Figure 5. In order to explain the map in the figure, the route from the point of departure to the destination and the objects on the map were described. In the figure, the triangular shape is the starting point, the rectangular shape is the destination, the fan shape are the landmarks, and the wedge shapes are the areas, the line segment is the route, and the diamond shapes are the nodes, respectively. Informational Text 3 is a text when turning around Node Q from the point at Node P and proceeding toward Node R. Informational Text 3 is as follows: "Turn right in front of the yellow ticket vending machine visible in the direction you are headed. After that, walk a little and proceed toward the map visible on the right. Here, the ticket vending machine and the map are selected as the landmarks for Nodes Q and R respectively. In Node Q, from the landmark position and the information about the wall, the ticket vending machine and the elevator are selected as visible landmarks. Next, the ticket vending machine, which is the largest among them, is chosen as a landmark. Likewise, in Node R, the map existing in the direction of movement is selected as a landmark.

With the method in 3.3, informational text is generated in the first half using Nodes P and Q and the ticket vending machine information. Here, from the positional relationships among Nodes P, Q and the ticket vending machine, " visible in the direction you are headed" is elicited, and using the features of the ticket vending machine, " yellow ticket vending machine" is given as supplemental information, and from the

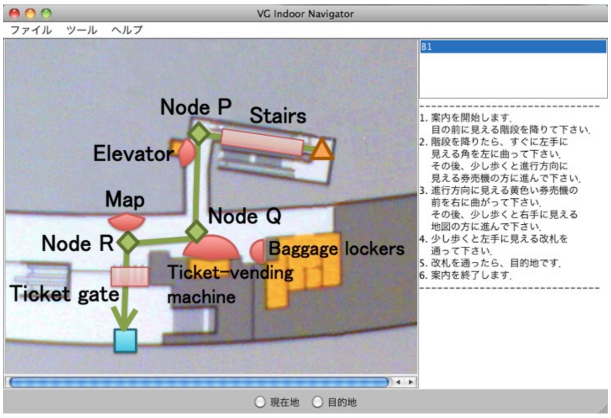

Figure 5: Informational text generation (Example 1)

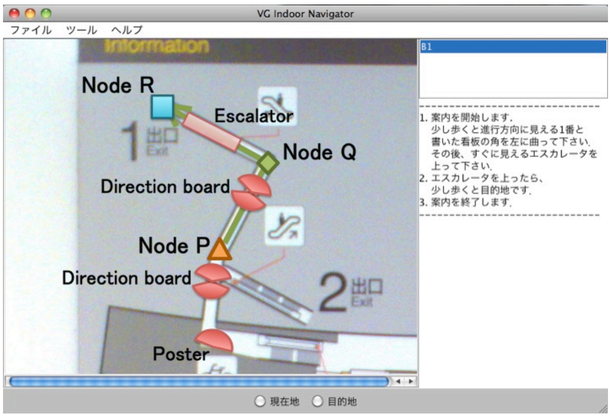

Figure 6: Iinformational text generation (Example 2)

positional relationship between Node Q and the ticket vending machine "Turn right in front of the ticket vending machine" is given. Also in the second half informational text is generated in the same way. Also, here the distance up to the landmark in the first half is not mentioned in order to abbreviate the informational text pertaining to distance, as it was deemed that it was at a distance directly visible from Node P, considering the information on the size of the ticket vending machine and the distance to the ticket vending machine from Node P.

By generating informational text in this manner, it was possible to provide directions on the next action using high-visibility landmarks. Also, the direction in which to proceed after the action was taken was also indicated with landmarks, and it was possible give directions on the direction in which to proceed after the action was taken.

We also discuss the generation of informational text examples in Figure 6 in a similar way. This is an informational text when proceeding to Nodes P, Q and R. Informational Text 1 provided directions as follows: " Walk a bit, and turn left at the corner with the billboard marked No. 1, which will be visible in the direction you are moving. Next, go up the elevator that will be immediately visible."

As landmarks existing near Node P, there are two billboards and a poster. However, in these two billboards they have different text written on their front and back sides. Therefore, a billboard is treated as two billboards, the respective text information is inputted, and the visible directions are registered from the respective directions they face. As a result, five landmarks exist near Node P: four billboards and one poster. From this, the landmarks existing in the direction of movement from Node P to Node Q are limited to one side of the billboard. Therefore, in the informational text, it is possible identify the text written on the billboard and use accurate supplemental information like "the billboard marked Number 1."

5 Method of Collecting Landmarks Used in Navigation

To make navigation possible, structural map information from various places is required. Structural map information is map information with structural information, such as walls, areas and rooms that is added to generally used map information. Micello Maps [7] and Google Sketch Up [8] , etc. are presented as examples of existing structural information in indoor spaces. Micello Maps are indoor maps of shopping malls, etc. created by Micello, Inc., and it only has information on specific facilities. Also, the data cannot be used freely. Google Sketch Up is a structural map for visualizing buildings, etc. in three-dimensions. Also, with the goal of providing an indoor position information service as an application of floor-based indoor positioning, we are engaged in research on the generation of indoor structural maps from a cloud sourcing approach. Even in outdoor spaces, with joint editing by many users with OpenStreetMap [9], a map that can be freely used is created.

Whether indoors or outdoors, landmarks which can serve as markers for specified places become the base point for the user of a position service to grasp their position and express it. For that reason, it has a major role in a diverse position information service. However, databases which have collected landmark information in detail do not exist. Numerous landmarks exist in the space targeted by location information services. Therefore, a system in which registration and correction of information by numerous users is necessary. We have successfully collected timetable information on the trains operating in Japan by cloud sourcing: wireless LAN position information with Locky.jp [10][11], and TimeTable.Locky [12]. Also, the collection of landmark information is possible using cloud sourcing, where users register landmarks on structural maps based on flat floor maps.

6 Conclusion and Future Issues

In this study, we proposed a method to make voice-navigation based on the visibility of landmarks possible. In the voice-navigation we proposed, using high-visibility landmarks which serve as position markers, the user's actions are guided. As clearer directions are possible at the point of action, it is effective with voice-navigation that does not have looking at a map as a premise. Also, in order to reduce route errors after the pedestrian takes an action, such as turning a corner, directions can be given using landmarks in the direction of movement after the action has been taken. We implemented a prototype system mounted with this method, verified its practicality, and finally examined the collection of landmark information for use in navigation.

As a future issue, the navigation system should estimate pedestrian location. In this study, it is difficult for location estimation to use GPS positioning because this system applies to indoor space. We intend to apply Scene Analysis to our current research [13]. In this paper, although we proposed giving directions using two-dimensional map information, it will be difficult to give directions on the landmarks and areas that will gradually come into view three-dimensionally. We were unable to consider situations where the pedestrian could not recognize the landmark, since it was obstructed by another landmarks, calculating the visibility of landmarks three-dimensionally will also be necessary. It may be useful to also change the timing of voice guides based on user's next action, the surrounding environment and circumstances. Therefore, we intend to explore the optimal timing based on the experiment.

REFERENCES

[1] NAVITIME. `http://www.navitime.co.jp/`.

[2] Google Maps Navigation(Beta). `http://www.google.co.jp/intl/en_us/mobile/navigation/`.

[3] Babaguchi, N., Dan, S., and Kitahashi, T. Generation of sketch map image and its instructions to support the understanding of geographical information. In *Pattern Recognition, 1996., Proceedings of the 13th International Conference*, Vol. 3, pp. 274 –278, 1996.

[4] Fujii, K. and Sugiyama, K. A Method of Generating a Spot-Guidance for Human Navigation. *The Journal of the Institute of Electronics, Information and Communication Engineers*, Vol. 82, No. 11, pp. 2026–2034, 1999.

[5] Tsukada, K. and Yasumura, M. ActiveBelt: Belt-Type Wearable Tactile Display for Directional Navigation. In *UbiComp 2004: Ubiquitous Computing*, Vol. 3205, pp. 384–399. 2004.

[6] Nakazawa, K., Kita, N., Takagi, K., Inoue, T., Shigeno, H., and Okada, K. A Dynamic Map Based on Landmark's Visibility(Geographical Information Services. *Transactions of Information Processing Society of Japan*, Vol. 49, No. 1, pp. 233–241, 2008.

[7] Micello Maps. `http://www.micello.com/`.

[8] Google SketchUp. `http://sketchup.google.com/`.

[9] OpenStreetMap. `http://www.openstreetmap.org/`.

[10] Yoshida, H., Ito, S., and Kawaguchi, N. Evaluation of Pre-Acquisition Methods for Position Estimation System using Wireless LAN. *The Third International Conference on Mobile Computing and Ubiquitous Networking*, pp. 148–155, 2006.

[11] Ito, S., Yoshida, H., and Kawaguchi, N. Studies on collection method of Access Point in metropolitan-scale 802.11 Location Systems. In *The Eighth International Conference on Ubiquitous Computing(Ubicomp2006)*, 2006.

[12] Yano, M., Kaji, K., and Kawaguchi, N. TimeTable.Locky: nation wide WiFi location information system based on user contributed information. In *Proceedings of the 3rd International Workshop on Location and the Web*, LocWeb '10, pp. 7:1–7:2, 2010.

[13] Ito, S. and Kawaguchi, N. Bayesian Based Location Estimation System Using Wireless LAN. *Proceedings of Third IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 273–278, 2005.