# A Proposal of IndoorGML Extended Data Model for Pedestrian-Oriented Voice Navigation System

Hirokazu Iida
Nagoya University, Japan
iida@ucl.nuee.nagoya-u.ac.jp

Kei Hiroi
Institute of Innovation for Future Society, Nagoya University, Japan
k.hiroi@ucl.nuee.nagoya-u.ac.jp

Katsuhiko Kaji
Faculty of Information Science, Aichi Institute of Technology, Japan
kaji@aitech.ac.jp

Nobuo Kawaguchi
Institute of Innovation for Future Society, Nagoya University, Japan
kawaguti@nagoya-u.jp

## ABSTRACT

We propose Landmark-Conscious Voice Navigation as one type of a pedestrian navigation system, which navigate users by only voice guidance. It is necessary to standardize data model in order to use this system widely. In a previous paper[1], we constructed a basic voice navigation system, which uses Open Street Map based data model. In this paper, at first, we conduct an experiment of voice navigation at an underground shopping area of Nagoya Station with two types of landmark descriptions. After that, we discuss what data structure is necessary to describe landmark information for voice navigation. Therefore, we propose to extend IndoorGML1.0 by adding landmark space as a new defined data model for voice navigation. The main contribution of this paper is that we conduct an experiment of voice navigation and research how different landmark descriptions affect users; furthermore, we discuss a IndoorGML extended data model for voice navigation.

## Categories and Subject Descriptors

H.2.1 [**Database Management**]: Logical Design—*Data models*; H.5.4 [**Information Interfaces and presentation**]: Hypertext/Hypermedia—*Navigation*

## General Terms

Design

## Keywords

IndoorGML, data model, landmark, pedestrian navigation, voice navigation


ISA'15, November 03-06 2015, Bellevue, WA, USA

ACM 978-1-4503-3978-0/15/11…$15.00
DOI: http://dx.doi.org/10.1145/2834812.2834814

## 1. INTRODUCTION

Recently, smart phone has became widely used device, and pedestrian navigation has became more popular. We have studied voice only navigation as one type of a pedestrian navigation system[1]. This system makes use of landmarks, such as shops, signs and stairs to generate voice guidance. Users need not to gaze at the display to check the navigation because this system provides only voice guidance. Thus, users can safely and smoothly walk their path. In indoor space, comparing with outdoor space, users can move freely on such as wide roads and halls. Thus, how to describe landmarks' features to the user is an important component for this navigation to be successful.

IndoorGML is an OGC standard for an open data model and XML schema for indoor spatial information[2]. It is defined as an application schema of GML, and it aims to provide a common framework of representation and exchange of indoor spatial information. But current version of IndoorGML1.0 cannot provide detailed information about landmarks, so we cannot use it for our voice navigation.

Standardizing a data model is valuable to applications such as voice navigation. For example, developers would not have to neither define data models nor create datasets by themselves. We can share datasets created along the standard data model and implement several types of applications.

We have constructed a basic pedestrian voice navigation system in Japanese that uses an OSM(Open Street Map [3]) based data model. This data model can generate voice guidance like, 'In 10 meters, please turn right at the T-junction where the red vending machine is on your left-hand side. After that, please move directly towards the elevator in front of you'(Figure 1). The important point of this guidance is that we use words like 'red vending machine' and 'elevator' as landmarks to help users navigate. Because of this, users can get sense of assurance and go their path smoothly.

In this paper, at first, we explain relevant study and problem about landmark based navigation and a data model for navigation. Next, we describe our basic voice navigation system. Using this basic system, we conduct an experiment of voice navigation with two types of landmark descriptions in order to research how these different landmark descriptions affect users. Then, we consider the data model for voice navigation, especially, on how to represent and describe landmarks. Finally, we propose an IndoorGML extended data model for pedestrian voice navigation by adding landmark definition space to IndoorGML.

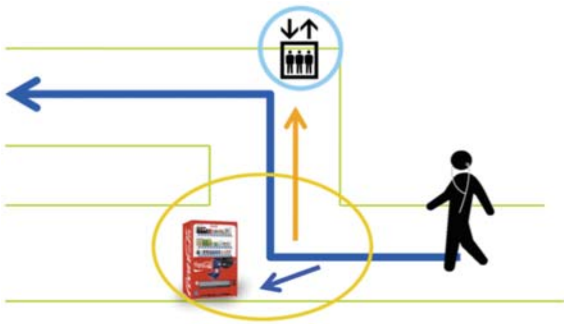


**Figure 1: Voice Navigation System**

# 2. LANDMARK BASED NAVIGATION AND STANDARD DATA MODEL

## 2.1 Landmark Based Pedestrian Navigation System

Nitta[4] and Fujii[5] have studied landmark based pedestrian navigation. Nitta's study was directed to realize of highly practical pedestrian navigation where no or partial positioning infrastructure is available. They proposed a method to estimate the current position from the interaction to check the visibility of each landmark as is done when a person guides the pedestrian via cell phone. Their method can provide the user a sense of conviction, however a map or text is displayed to help the user navigate, so it may cause danger due to the user's lack of attention.

Fujii used space-object landmarks for their navigation. These types of landmarks are figures shown in maps such as a building or a park. In order to improve navigation more intelligibly, it is necessary to use not only such space-object landmarks, but also, signs or more landmarks' features such as color or height(i.e: 8-meter tall golden clock).

## 2.2 Landmark-Conscious Pedestrian Voice Navigation

We have proposed voice only pedestrian navigation, which can navigate users without gazing the display. Because our navigation system provides only voice guidance, users need not to stop and look at a map on display and therefore they can walk safely. We use highly visible objects such as shops, signs and stairs as landmarks in order to navigate intelligibly. In order to navigate users by using only voice guidance with visible landmark, it is necessary that the system can compute visible landmarks and select the highest visible landmarks to the user. And it is also necessary that the system considers how to describe the landmarks' features in order to convey instructions by only voice guidance.

## 2.3 Standard Data Model

If a system uses datasets that were created along system's original data model, data sets can be used only by this particular system. But if each system accepts the same standard data model, they can share datasets and several navigation systems would be available using the same datasets. In addition, datasets maintenance costs will be cut down by managing them as a whole.

## 2.4 IndoorGML

IndoorGML[2] is a standard data model created by OGC (Open Geospatial Consortium[6]), which consists of Core Module and Indoor Navigation Module. Indoor space is defined as space within one or multiple buildings consisting of architectural components in IndoorGML. An indoor space is considered as a set of cells, which are defined as the smallest organizational or structural unit of indoor space. And each cell has information to represent semantics, geometry and topology.

IndoorGML supports multiple representation layers with different cellular spaces for an indoor space(Figure 2). Each semantic interpretation layer results in a different decomposition of the same indoor space where each decomposition forms a separate layer of cellular space.

However, IndorGML1.0 has not defined detailed information about landmarks. Thus, we cannot apply it to voice only navigation.

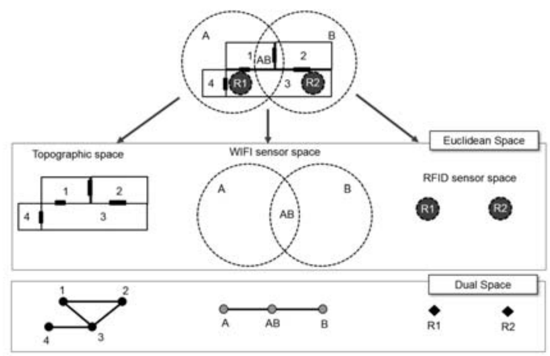



**Figure 2: Multi-Layered Representation[2]**

# 3. DESIGN OF BASIC PEDESTRIAN VOICE NAVIGATION SYSTEM

We have designed a basic pedestrian voice navigation system that uses OSM based original data model. In this section, we describe about this system and the data model. Then, we discuss landmark information for voice navigation.

## 3.1 Structure of Basic Voice Navigation System

Figure 3 shows system structure of our basic voice navigation system. First, the system gets departure place and place of destination from the user. Then, the system generates route that the user should pass, and decides the point where the system conveys guidance such as junction, stair and etc. After that, the system generates contents of voice guidance for each guidance point. At this stage, it is important how to select the highest visible landmark and how to describe the landmark's feature. Voice guidance is automatically conveyed to the user when the user arrive each guidance point.

The information that the system requires is shown below.

- Space Structure Information
  - Building structure
  - Pedestrian network
- Landmark Information
  - Information to compute visibility of the landmark
  - Information to describe the landmark

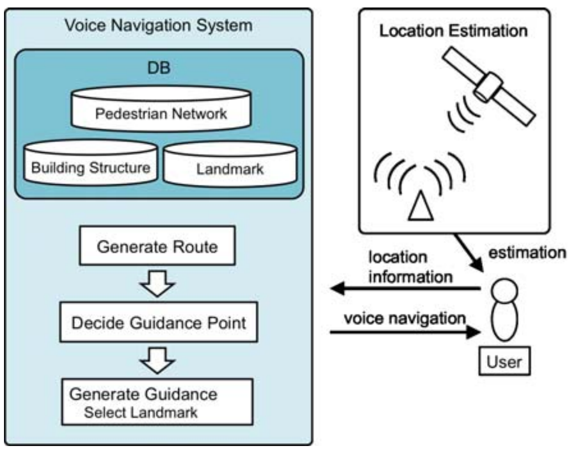



**Figure 3: Structure of Basic Voice Navigation System**

## 3.2 OSM Based Original Data Model

We defined data model to represent data structure as shown in section 3.1. It consists of three parts, Node class, Way class and Relation class. Each class has tags to represent semantics such as building structure, pedestrian network, intersections, stairs, landmarks and etc. Each tag has a key and a value fields to represent several peace of information.

- Node class
  - Represents coordinates.
- Way class
  - Represents set of nodes to generate objects and network.
- Relation class
  - Represents set of object or network.

Landmark information is defined by further detailed tags as shown in Table 1. Landmark is represented as Way class and the system can calculate size of the landmark and direction from the user's line of sight.

**Table 1: Example of Landmark Information Tags**

| Key | Value | explanation |
|---|---|---|
| amenity | toilets,atm | type of facilities |
| shop | clothes,shoes | genre of shop |
| height | ceiling,eye level | height of the landmark |
| color | red,blue | color of the landmark |
| destination | 'Exit 1' | text written in sign |
| pronunciation | 'DIANA' | proper noun of the landmark |

## 3.3 Landmark Information for Voice Navigation

Because the system do not use display to describe landmark information, it is important to consider how to create landmark description for representing it by only voice guidance. Hansen[7] proposed a classification of landmarks according to their function in route guidance. In this part, we explain the process of utilizing landmark in voice navigation and information for it. (1)Detect visible landmark from the user (2)Select the highest visible landmark (3)Generate landmark description.

### *3.3.1 Detect visible landmark*

At first, the system detect all visible landmarks from the user's location. For detecting visible landmarks, space structure that connects visible space is necessary. Using this space structure, all the visible landmark objects can be used even if the user see a landmark across such as a glass partitions, wellhole and etc. We should also consider time availability of the landmark. For example, some shops may pull down a shutter off business hours and a user won't be able to recognize it. Some signs may be not able to be visible with a certain lighting situation.

It is also necessary to consider from which direction the landmark can be recognized. For example, some sign landmarks are recognized from one side or have different information on each side.

### *3.3.2 Select the highest visible landmark*

It is necessary that the system can select the highest visible landmark from all visible landmarks. It depends on positional relation between the landmark and the user and landmark feature such as color, written text, size, height and etc. It is also affected by surround environment of the landmark and user's profile information.

### *3.3.3 Generate landmark description*

Description of landmarks is also important factor to smoothly navigate users by using only voice guidance. According to landmark description, how cognitive the landmark is will be affected even though the system uses the same landmark. Landmarks are described by using landmark name with direction from the user and landmark feature(color, written text and etc.). This is affected by other elements such as positional relation between the user and the landmark, environment conditions and user's profile information, just like selecting the highest visible landmark.

# 4. EXPERIMENT AT AN UNDERGROUND SHOPPING AREA OF NAGOYA STATION

How to describe the landmark in voice guidance is the important element in order to navigate intelligibly. In order to research how different landmark descriptions affect users, we conducted an experiment of voice navigation with two types of landmark descriptions. One type of landmark description is proper noun(shop name), the other is common noun(genre). We conducted the experiment at an underground shopping area of Nagoya Station where there are mainly fashion shops(clothing, accessories and etc.) for women. Obviously, it is difficult to navigate users with common noun described landmarks because some other objects are in the same genre.

## 4.1 Outline of the Experiment

We made datasets of the underground shopping area of Nagoya Station, which contains landmark's information consisted of shops, signs and stairs. Figure 4 shows examples of landmark information. Each landmark possesses a pronunciation tag to describe it with a proper noun and shop tag to describe it with a common noun. Table 2 shows examples of landmark described by proper noun and common noun. Sign landmarks have height tag and can be used for users to identify the landmark more quickly. Using this kind of information, we automatically generate voice guidance like, 'Please turn right when you can see DIANA on your left-hand side. In a few minutes, you will find a junction where AMO'S STYLE is on your right-hand side'. Voice guidance conducts at each point where users need to do some actions such as turning, taking the stair or passing through intersection. If there are some visible landmarks, we use the landmark that is closest to the user's line of sight. We used Pedestrian dead reckoning with magnetic field and WiFi fingerprints[8] as location estimation in this experiment. Normally, voice guidance is conveyed automatically, but if the guidance is not conveyed caused of location estimation error, users can listen the guidance by pressing a button attached to the earphone.

The experiment was participated in by students in their 20's including 9 men and 1 woman, and navigated four routes for each subject. Each subject was navigated two routes by voice guidance with proper noun landmark descriptions and two routes with common noun landmark descriptions. Figure 5 shows the example of navigated route and voice guidance. After the navigation, we investigated their navigation experiences with open-ended questionnaires.

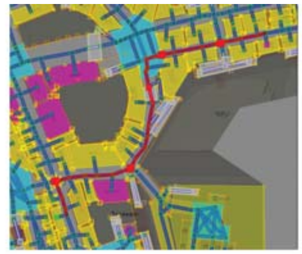

```
<way id="4167">
  <tag k="name" v="SUDEST ユニモール"/>
  <tag k="poi" v="1465"/>
  <tag k="pronunciation" v="シュッドエスト"/>
  <tag k="reference point" v="yes"/>
  <tag k="shop" v="clothes"/>
  <nd ref="4165"/>
  <nd ref="4166"/>
  <nd ref="4168"/>
  <nd ref="4169"/>
  <nd ref="4165"/>
</way>
```

```
<way id="46313">
  <tag k="destination" v="近鉄、名鉄"/>
  <tag k="destination sign" v="yes"/>
  <tag k="height" v="2.5"/>
  <tag k="name" v="構内案内板"/>
  <tag k="pronunciation" v="構内案内板"/>
  <tag k="reference point" v="yes"/>
  <tag k="transparent" v="yes"/>
  <nd ref="46311"/>
  <nd ref="46312"/>
</way>
```


Figure 4: Example of Landmark Data

Table 2: Example of Landmark Description

| proper noun | common noun |
|---|---|
| SUDEST | clothing store |
| DIANA | shoe store |
| direction board for KINTETSU, MEITETSU | sign |



**By Proper Noun**

- Let's start! Please move directly towards PROPORTION BODY DRESSING.
  In a few minutes, you will find a junction.
- Please keep going straight once you can see the junction
  In a few minutes, you will find a hall where UNIVERVAL MUSE is on your right-hand side.
- Please turn left when you can see the hall
  where UNIVERVAL MUSE is on your right-hand side.
  In a few minutes, you will find stairs.
- Please take the stairs when you can see stairs.
  In a few minutes, you will find a junction where
  a sign board that says MIDLAND is on your overhead left-hand side.
- Please turn right when you can see the junction
  where the sign board that says MIDLAND is on your overhead left-hand side.
  In a few minutes, you will find a junction
  where a direction board that says KINTETSU is overhead.
- Please turn left when you can see the junction
  where the direction board that says KINTETSU is overhead.
  In a few minutes, you will arrive at your place of destination.
- You arrived at your place of destination. End of navigation.

--------------

**By Common Noun**

- Let's start! Please move directly towards clothing shop.
  In a few minutes, you will find a junction.
- Please keep going straight once you can see the junction
  In a few minutes, you will find a hall
  where clothing shop is on your right-hand side.
- Please turn left when you can see the hall
  where clothing shop is on your right-hand side.
  In a few minutes, you will find stairs.
- Please take the stairs, when you can see stairs.
  In a few minutes, you will find a junction
  where a sign is on your overhead left-hand side.
- Please turn right when you can see the junction
  where the sign is on your overhead left-hand side.
  In a few minutes, you will find a junction where a sign is overhead.
- Please turn left when you can see the junction where the sign is overhead.
  In a few minutes, you will arrive at your place of destination.
- You arrived at your place of destination. End of navigation.

Figure 5: Example of Guidance

## 4.2 Examination

In this experiment 80 percents of them could arrive at their destinations. This means that the feasibility of voice navigation that uses auto-generated voice guidance using our datasets. The main reason why they failed was that location estimation did not work well and they also forgot to press the button for guidance. Thus, these subjects were not able to find their way. The other reason is that they mistakenly took another landmark instead of the described landmark that was described with the same common noun.

We got valuable results from the questionnaires. In the case that a landmark was described with common noun, the subject misunderstood when the system described an underwear shop as 'clothing shop', and also, when the system described a clothing shop which handles not only clothing but also shoes, accessories. There were some shops that looked similar to other shops nearby. Thus, subjects sometimes could not identify the landmark from common noun description. However, when subject could identify a landmark, the common noun description is simple and easy to hear.

In the case that landmark was described with proper noun, the subject sometimes could not catch complicated landmark names, such as 'PROPORTION BODY DRESSING'. Also a subject could not find a shop that had a sign in a language other than Japanese or English(Figure 6) because he could not connect described landmark to shop's sign in real world. However, except of such cases, proper noun description won't cause misleading and users can easily identify described landmark.

We can also find from the experiment's result that the navigation result changes drastically for different types of landmark descriptions(proper noun vs. common noun). And, the data structure should represent information to generate several landmark descriptions depending on landmark surrounded environment and user's demands.

In this experiment, we selected a landmark that is closest to user's line of sight as the most visible landmark. However, we have to consider how to select the landmark that is the easiest to be identify in real world environment for intelligible navigation.


**Figure 6: Sign Which is Difficult to Read for Japanese**

## 5. INDOORGML EXTENDED DATA MODEL

It is necessary to standardize the data model in order to use voice navigation system widely. IndoorGML can represent space structure information, however it cannot represent detailed landmark information for voice only navigation. In this section, we arrange information about landmarks for voice navigation by considering the experiment results. Then, we propose IndoorGML extended data model by adding definition about landmarks and how to make the best use of it.

### 5.1 Definition of Landmark Information

We used shops and signs as landmarks in the experiment. However, there are many kinds of landmarks such as vending machine, ATM, monument, crossroads, etc. Thus, we divide landmarks into categories in order to represent these landmarks as shown in Table 3.

Because each landmark has different features, each landmark category is individually defined with more detailed particular information. We define FacilityCategory and SignCategory at first from the results of the experiment.

FacilityCategory categorizes facilities such as clothing shop, coffee shop and etc. It can define genre of landmark, proper noun and handled item. It can also define sign that contains written text, text color, color, appearance if the landmark has sign.

**Table 3: LandmarkCategory**

| category | example |
|---|---|
| facility | closing shop, coffee shop |
| sign | direction board, advertisement |
| equipment | vending machine, ATM |
| art | monument, art |
| topography | stairs, crossroads |

SignCategory categorizes signs such as direction board, advertisement and etc. It can define color, appearance and guidance information. Guidance information describes information of direction board. Phan[9] also mentioned that this kind of information is important to navigation.

Actually, there are still other LandmarkCategorys to represent landmarks in indoor space. We would like to consider more about this information as a future work.

### 5.2 Structure of IndoorGML Extended Data Model

IndoorGML is a standard data model and focus on indoor spaces for navigation purposes. Because it can represent space structure as not only 2D but also 3D model, it is suited to calculate visibility of landmark. And it can contain information for location estimation. Current version of IndoorGML can define building structure, pedestrian network and sensor information for location estimation. However, it has not defined detailed landmark information for voice navigation. Thus, we propose IndoorGML extended data model for voice navigation by adding landmark space layer to IndoorGML. Figure 7 shows structure of proposed data model.

Landmark information is defined in LandmarkSpaceLayer. Each landmark has to contain minimum definition, at least one description. If you would like to define more information, you can define it as LandmarkCategory, mentioned in 5.2. It is considered that navigation system can compute visibility of the landmark and generate several descriptions by using these pieces of information.

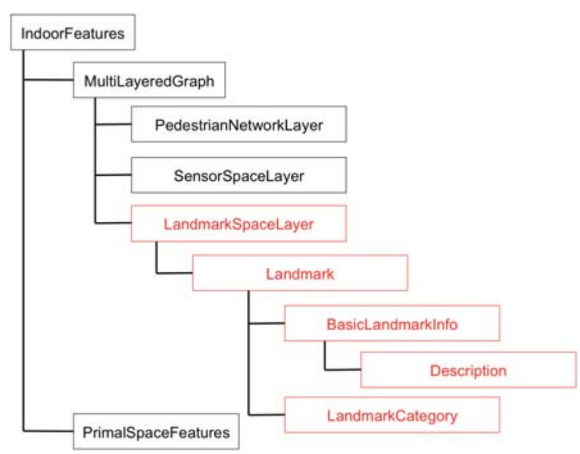



**Figure 7: Structure of Proposed Data Model**

## 6. CONCLUSION

In this study, we proposed IndoorGML extended data model to define voice navigation standard data model. At first, we implemented our basic voice navigation system and conducted an experiment at an underground shopping area of Nagoya Station in order to research how different landmark descriptions affect users. After that, we considered how to represent landmark information for voice navigation. And finally, we defined landmark space layer in order to use landmarks in navigation.

As a future improvement, we have to define more detailed information to each LandmarkCategory and to consider how to compute visibility of the landmark and what information is necessary for it. Also, we have to consider methods to collect landmark information and widely provide our navigation system. Kyriakos' study[10] focused on crowdsourcing to provide indoor information service and investigating it into IndoorGML. Biczok[11] used indoor positioning and location-based service, called MazeMap, in a university campus and succeeded to collect valuable data.

## 7. REFERENCES


[1] Shota Watanabe, Katsuhiko Kaji, and Nobuo Kawaguchi. A proposal of landmark-conscious voice navigation. *International Conference on Mobile Computing and Ubiquitous Networking*, pp. 56–61, May 2012.

[2] IndoorGML.http://docs.opengeospatial.org/is/14-005r3/14-005r3.html. Online: accessed 10-Sep-2015.

[3] OpenStreetMap.https://www.openstreetmap.org. Online: accessed 10-Sep-2015.

[4] Tomoyuki Nitta, Kazuya Miyazaki, Shun Yoshimi, Ryosuke Tabata, Ismail Arai, Takuya Azumi, and Nobuhiko Nishio. Visibility confirmation based pedestrian navigation system. *Information Processing Society of Japan Interaction*, Vol. 2EBX-27, pp. 448–455, March 2013.

[5] Kensaku Fujii and Kazuhiro Sugiyama. A method of generating a spot-guidance for human navigation. *The transactions of the Institute of Electronics, Information and Communication Engineers*, Vol. J82D, No. 11, pp. 2026–2034, November 1999.

[6] Open Geospatial Consortium.http://www.opengeospatial.org/. Online: accessed 10-Sep-2015.

[7] Stefan Hansen, Kai-Florian Richter, and Alexander Klippel. Landmarks in openls - a data structure for cognitive ergonomic route directions. *Geographic Information Science 4th International Conference, GIScience 2006, Munster, Germany*, pp. 128–144, 2006.

[8] Ryoji Ban, Katsuhiko Kaji, Kei Hiroi, and Nobuo Kawaguchi. Indoor positioning method integrating pedestrian dead reckoning with magnetic field and wifi fingerprints. *The Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU2015)*, pp. 169–174, January 2015.

[9] Phan Duy Phuoc, Tomohiro Miyajima, Akihiro Birou, Takuya Azumi, and Nobuhiko Nishio. Pedestrian indoor navigation methods for non-corridor area traverse. *Information Processing Society of Japan Technical Report*, Vol. 2013-UBI-38, No. 56, pp. 1–6, May 2013.

[10] Kyriakos Georgiou, Timotheos Constambeys, Christos Laoudias, Lambros Petrou, Georgios Chatzimilioudis, and Demetrios Zeinalipour-Yazti. Anyplace: A crowdsourced indoor information service. *Proceedings of the 16th IEEE International Conference on Mobile Data Management*, Vol. 2, pp. 291–294, 2015.

[11] Gergely Biczok, Diez Martinez, Thomas Jelle, and John Krogstie. Navigating mazemap: Indoor human mobility, spatio-logical ties and future potential. *Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pp. 266–271, March 2014.