

LandmarkSense: A Mobile Sensing System for Automatic Detection of Railway Stations Landmarks

MOUSTAFA ELHAMSHARY, AKIRA UCHIYAMA, HIROZUMI YAMAGUCHI, TERUO HIGASHINO^{1,a)}

Abstract:

We present *LandmarkSense*, a novel mobile sensing system for precisely detecting various landmarks (e.g., ticket vending machines, entrance gates, drink vending machines, lockers, etc) that exist in railway stations. Our key observations show that certain locations (i.e., landmark locations) in railway stations present identifiable signatures on one or more cell-phone sensors. A ticket vending machine, for instance, imposes a distinct pattern on a smartphone's accelerometer and gyroscope as well as it experiences an unusual magnetic fluctuation. *LandmarkSense* leverages this fact to automatically recognize these landmarks to enable a myriad of travel support applications. We evaluate *LandmarkSense* through a field experiment in a major train and subway stations in Japan. Our results show that *LandmarkSense* can detect different landmarks accurately with at most 9.7% false positive rate and 7.4% false negative rate for all types of landmarks. Moreover, we show that *LandmarkSense* has a small energy footprint on cell-phones, highlighting its promise as a ubiquitous travel support service.

Keywords: Railway Stations, Activity Recognition, Floorplan Construction

1. Introduction

With the fact that people spend most of their time at indoor spaces, indoor Location Based Services (LBSs) are being developed at a phenomenal rate with a variety of applications including mapping and navigation services, point-of-interest finders, geo-social networks, and advertisements. A key requirement to indoor LBSs is the availability of indoor maps to display the user location on. Realizing the economic value of this technology, a number of commercial navigation systems for indoor mapping have started to emerge. In late 2011, Google Maps started to expand its coverage by providing detailed floorplans for a few malls and airports in the U.S. and Japan as well as allowing buildings owners around the world to upload their indoor floorplans. Nevertheless, these maps are still limited in coverage to a small number of countries featuring only some major airports, shopping malls, etc. This limitation in coverage is due in part to the following reasons: (1) buildings owners may not allow sharing of their floorplans in public for privacy reason, (2) buildings internal structures often evolve over time, furthermore (3) manual creation of these maps requires slow, labor-intensive tasks, and they are subject to intentional incorrect data entry by malicious users.

Railway stations, as an example of indoor places, are a key part of the day-to-day lives of people having millions of passengers every day (e.g., Shinjuku station in Japan has 3.64 million passengers/day on average in 2007). In highly populated countries, major stations have large indoor spaces (e.g., Shinjuku station has 36 platforms and over 200 exits). Furthermore, the coverage ratio

of railway stations by commercial navigation systems is still limited, with Google indoor maps covering less than 50 transit stations worldwide which are only a small fraction from thousands of stations on the earth. The lack of detailed digital floorplans for stations highlighting locations of various landmarks worsens passengers' experience, especially foreigners or first-time visitors. Consequently, this sparks the need for the automatic construction of detailed indoor floorplans for railway stations.

To resolve this problem, the research community recently has embarked to address the problem of automatic construction of indoor floorplans by exploiting motion trajectories of mobile phone users [1], [2], [3]. These systems proved the feasibility of estimating the general layout of a building [1], [2], [3], identifying rooms shape and dimensions [1], [3], along with identifying other points of interest such as store entrances [1], [3]. Nevertheless, none of these approaches provide semantic-rich floorplans where various landmarks are tagged on the floorplan that are necessary for many of today's map-based applications. For example, stations indoor navigation systems should rely on important landmarks to better guide passengers to their destinations; station evacuation planning is ineffective if maps are not tagged with stairs used as emergency exits; and a person with disability needs a map that shows elevator-enabled routes.

In this paper, we present *LandmarkSense* as a crowdsensing system that leverages the ubiquitous sensors available in commodity cell-phones to automatically enrich railway stations' floorplans with different landmarks. These landmarks are essentially certain structures in the building -stairs, elevators, escalators- or station installed machines -vending machines, ticket gates, etc- that force users to behave in predictable ways. These predictable behaviors can be translated to sensor signa-

¹ Grad. Sch. of Info. Sci. & Tech., Osaka University

^{a)} m.elhamshary, uchiyama, h-yamagu, higashino@ist.osaka-u.ac.jp

tures. For instance, a passenger crossing an entrance gate has to slow down her walking speed until she pauses to drop the ticket into the gate machine and then steps forward to grab it. Meanwhile her phone is experiencing a magnetic field distortion emanating from the gate machine electronics. Therefore, starting from an **unlabeled general** floorplan of a station, *LandmarkSense* will be able to estimate the location of different station landmarks and tag their locations on the map accordingly to generate a detailed floorplan.

Translating this basic idea into a deployable system, however, involves addressing a number of challenges: first, identifying landmarks signatures from the sensed data warrants unsupervised learning on sensor features. Second, *LandmarkSense* leverages the pedestrian dead-reckoning (PDR) technique to estimate passengers' location at the time of activities. Since PDR has an average localization error in the range of few meters [4], it can place the passenger in a location on the floorplan that is far from the actual one. Finally, the system needs to be optimized for energy to avoid a significant battery drain.

LandmarkSense's design addresses these challenges and its implementation over different Android phones shows that it can detect different station landmarks accurately with at most 9.7% false positive rate and 7.4% false negative rate for all types of landmarks. This comes with a low power consumption of 46mW on the average.

In summary, our contributions are three-fold:

- We present the *LandmarkSense* system to automatically crowdsense and identify stations landmarks from phone sensors without imposing any overhead on the passenger and with minimal energy consumption.
- We provide a framework for extracting different features from phone sensors to identify different stations' landmarks.
- We have collected real data by 9 participants, implemented *LandmarkSense* on Android phones, and evaluated its accuracy and energy-efficiency in major subway and train stations in Osaka.

2. System overview

Figure 1 shows the system architecture based on a crowdsensing approach, where cell phones carried by users submit their data to the server in the cloud. The data is first preprocessed to reduce the noise. Then, landmarks are classified to separate transport mode landmarks (elevators, escalator, and stairs) from other stations specific landmarks (ticket vending machines, entrance gates, etc). *LandmarkSense* has two core components: one for extracting transport mode landmarks and the other for extracting other stations specific landmarks. We take a classifier approach to detect the different landmarks based on the extracted features from the collected data. We give an overview of the architecture in the following subsections and leave the details for the landmarks detection to sections 3 and 4.

2.1 Traces Collection

The system collects time-stamped sensor measurements that include available inertial sensors, barometer as well as the sound sensor. Inertial sensors have a low cost energy profile and they

are already running all the time during the standard phone operation to detect phone orientation changes. Therefore, they consume zero extra energy. On the other hand, as the sound sensor consumes a little extra energy, we avoid continuous sensing of it by using an adaptive sensor scheduling scheme called triggered sensing[5]. The key idea is that not all sensors are sampled continuously where sensors that are inexpensive in energy consumption (e.g., accelerometer) are used to trigger the operation of more expensive sensors (e.g. sound). Specifically, *LandmarkSense* activates audio recording only as soon as the passenger becomes stationary and once she resumes walking again, audio recording is suspended. The intuition is that passengers traces are dominated by walking periods and they pause only to perform activities (e.g., buying tickets) which is associated with a landmark (e.g., ticket vending machine). The collected audio recording at activity time (i.e., using a landmark) is used as a tie breaker when other sensors (e.g. inertial sensor) fail to recognize this activity and thus identify its uniquely associated landmark.

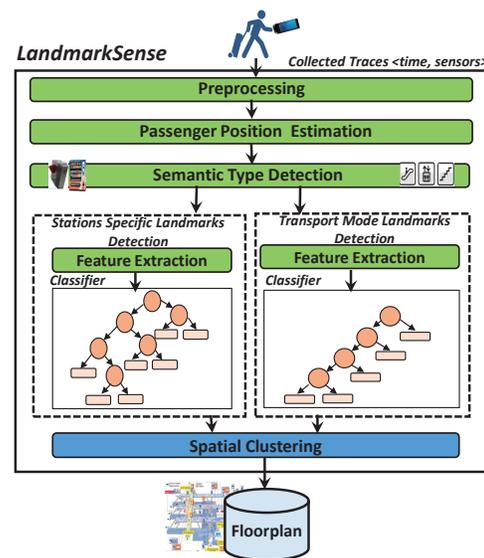


Fig. 1: The *LandmarkSense* system architecture.

2.2 Preprocessing

This module is responsible for preprocessing the raw inertial sensor measurements to reduce the effect of (a) phone orientation changes and (b) noise and bogus changes, e.g. sudden breaks, or small changes in the direction while moving. To handle the former, we transform the sensor readings from the mobile coordinate system to the world coordinate system leveraging the inertial sensors. To address the latter, we apply a low-pass filter to the raw sensors data using local weighted regression to smooth the data [6]. To filter out the noise in the employed frequency bands in audio recordings, the standard sliding window averaging technique is used.

2.3 Passengers Position Estimation

LandmarkSense needs accurate passengers locations during the usage of station landmarks to estimate landmarks positions. To

achieve this, *LandmarkSense* employs the dead-reckoning technique to track the passenger's location starting from a reference point (e.g., the station entrance). However, the displacement error of dead reckoning is unbounded making it infeasible for indoor tracking. To alleviate this problem, *LandmarkSense* incorporates the idea of the Unloc [7] by leveraging **amble and unique** physical points in the stations (i.e., landmarks) to reset the accumulated error. Since dead-reckoning provides a rough location to the phone, it is also possible to roughly localize the landmarks based on when the phone senses them. Now, since the floorplan is known, we can estimate the locations of all landmarks in a crowd-sensing approach (as discussed later) by combining the rough estimates (i.e., the dead-reckoned positions) from multiple passengers' phones. These landmarks, once detected based on their unique sensor signatures, can then be used to improve dead-reckoning of subsequent passengers, which in turn can refine the landmarks locations. This recursive dependence between estimating the landmark location and the user location is similar to the Simultaneous Localization And Mapping (SLAM) framework.

2.4 Features Extraction

LandmarkSense extracts numerous features from passengers' phone sensors to identify stations landmarks. Since some features are used to identify several landmarks, we will elaborate how these features are extracted to avoid redundancy. For instance, magnetic peak is key feature to recognize many landmarks that involve direct interaction with electronic machines (e.g., vending machines) or even passing through them (e.g., entrance gates). To extract this magnetic peak, we identify it by first applying a stream-based event detection algorithm to identify significant changes in the magnetic field. Once a significant change has been observed, we mark the corresponding time instant as the starting boundary of the peak area. We buffer subsequent measurements until a significant decrease in the magnitude of the magnetic field is observed. Once the starting and ending boundaries have been identified, we extract two features that characterize the peak area such as the peak period and its strength. Moreover, many activities are characterized by a sudden change in the user direction (i.e., surge in gyroscope readings) during or directly after the activity period. To detect this sudden change, we used the *approximate derivative* method. The derivative of sensor values within a time window are compared against a predetermined threshold to detect the surge in sensor values. Finally, the variance of the acceleration is used to discriminate various motion types of passenger (stationary, slow walking and normal walking) which are essential patterns that contributes to identify many activities. It should be noted that various thresholds have been used to identify landmarks from the features extracted from sensor data. These thresholds are determined empirically from the data collected during a preliminary experiment.

2.5 Landmark Type Detection

LandmarkSense is designed to detect various station landmarks based on their unique usage pattern. First, we separate the two major types of landmarks: transport mode landmarks (elevators,

stairs and escalators) and stations exclusive landmarks (e.g., ticket vending machines, entrance gates, etc). The usage of transport mode landmarks involves a noticeable change in the passenger's level (i.e., height) which is absent in other landmarks (Figure 3). To separate them, we draw on the maximum difference among the relative barometer readings (i.e., pressure) in consecutive overlapping windows. The intuition is that a change in pressure means a change in height which in turn means that the passenger is using one of the transport mode landmarks. Moreover, the sign of the pressure difference indicates the direction of motion (up or down) which is useful for other purposes (e.g., escalators direction). Later, the major two classes of landmarks are further classified to their more fine-grained landmarks.

2.6 Landmarks Extraction

To identify landmarks, *LandmarkSense* relies on a tree-based classifier to identify different landmarks as it is easy to understand and to generate its rules. Moreover, given that many landmarks share some sensors patterns while having different patterns on other sensors, the hierarchical classification (e.g., decision tree) is the optimized solution. The tree-based classifier decomposes the task hierarchically into subtasks, proceeding from a coarse-grained classification (shared patterns) towards the distinction of fine-grained landmarks (distinctive patterns) as detailed in Section 3 (Transport mode landmarks) and Section 4 (Stations specific landmarks).

2.7 Landmark Location Estimation

Whenever a landmark is detected by the landmark detection modules, *LandmarkSense* needs to determine whether it is a new instance of a landmark or not as well as determine its location. To do this, *LandmarkSense* applies spatial clustering for each type of the extracted landmarks. It uses the density-based clustering algorithm (DBSCAN [8]). DBSCAN has several advantages as the number of clusters is not required before carrying out clustering; the detected clusters can be represented in an arbitrary shape; and outliers can be detected. The resulting clusters represent samples of the discovered landmarks. After clusters are formed, the locations of the newly discovered landmarks are estimated as the weighted mean of the points inside their clusters. We weight the different locations based on their location accuracy reported by our position estimation: In our Unloc based location estimation approach, the longer the user trace from the last resetting point, the higher the error in the trace [7]. Therefore, shorter traces have better accuracy. Based on the law of large numbers, the weighted average of independent noisy samples should converge to the actual location of the landmark. When a new landmark is discovered, if there is a discovered landmark within its neighborhood, we add it to the cluster and update its location. Otherwise, a new cluster is created to represent the new landmark. To reduce outliers, a landmark is not physically added to the floorplan until the cluster size reaches a certain threshold which is specified by *minpts* parameter (the minimum number of points that can form a cluster) of the DBSCAN algorithm.

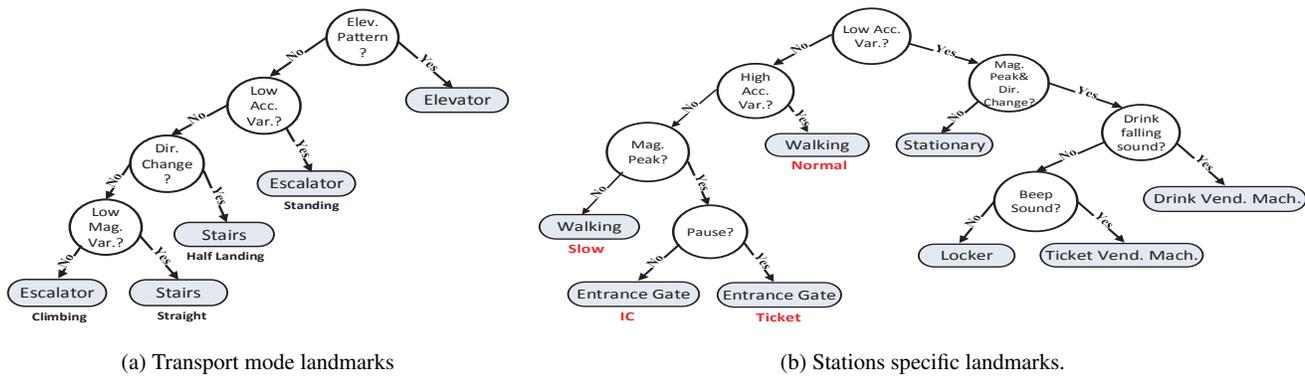


Fig. 2: A decision tree classifier for detecting different types of landmarks.

3. Transport Mode Landmarks Detection

This class of landmarks are based on using the inertial sensors which have the advantage of being ubiquitously installed on a large class of smart phones, having a low-energy footprint.

Elevator: We begin by separating elevator from the other classes as it is straightforward to distinguish its unique pattern. The typical usage scenario of an elevator consists of a normal working period, waiting for the elevator, walking into the elevator, changing direction to face the exit door, standing for a while, followed by a change in the level when it starts to move (Figure 4). This behavior is reflected to a unique pattern that consists of sequence of states: walking, stationary, stepping, direction change, level change, and Accelerate -Stationary-Decelerate emerging from the start and stop of the elevator. This multi-modal pattern is detected by using Finite State Machine (FSM) that receive the extracted features from accelerometer, gyroscope and barometer readings. It then depends on the observed state transitions where different thresholds are used to move between the states to recognize the elevator.

Escalator (Standing): The variance of acceleration can be used to separate some samples of escalators (when user is standing) from stairs. If the variance of acceleration is very low, it can be affirmative that the user is standing on the escalator since climbing stairs will absolutely generate a high acceleration variance that due to the vertical motion of user. On contrast, if the variance of acceleration is high, we are uncertain whether the user is using a stair or an escalator (e.g. some users climb escalators).

Half Landing Stairs: There are two types of stairs: straight and half landing stairs. Half landing stairs have a turn in their middle forcing the passenger to change her direction while straight ones do not have any turns. Thus, if there is a change in the user direction (i.e., a surge in gyroscope readings) in the middle of the level change period, it is an affirmative that the user is climbing a half landing stairs as all escalators are straight without any turns. On the other hand, if there is no change in user direction, we cannot verify the mode of transport given that some stairs are also straight.

Escalator (Climbing): Now we have to differentiate between climbing a straight stair and climbing an escalator after separating all other traces. To separate them, we found that the variance

of the magnetic field that due to the escalator machinery can be a reliable discriminator as shown in Figure 5.

Straight Stairs: After separating other transport mode landmarks, the remaining samples are classified as straight stairs.

Level Change or Floor Change: Many stations are multi-floor buildings with a typical floor height between 3.0 to 6.0 meters. The majority of transport mode landmarks move passengers from one floor to another (Floor change landmarks). However, there exist some low height stairs and escalators which move passenger from level to another within the same floor (level change landmarks). To classify the type of escalators and stairs (marked by red stars in Figure 2a), we rely on the magnitude of pressure difference during the elevation change period. Given that 1.0 meter height change corresponds to 0.12 hPa change in pressure, the pressure difference of 0.3 hPa is used as a threshold to separate level change escalators and stairs from floor change ones.

4. Stations Specific Landmarks Detection:

Stations are rich with many exclusive landmarks like ticket vending machines, entrance gates, drink vending machines and lockers. In the next subsections, we give the details of how to identify these landmarks.

4.1 Coin Operated Machines

To identify coin operated machines such as ticket and drink vending machines and luggage lockers, we observed that their typical usage traces consist of normal walking to the machine, followed by standing in front of it, inserting currency, beginning the service (choosing a drink or the ticket type in case of drink and ticket vending machines respectively or opening the locker door in case of lockers), finishing the service (grabbing the ticket or the drink in case of ticket and drink vending machines respectively or putting luggage into the drawer and locking it), and finally walking away (Figure 6). This usage scenario is translated to the following unique patterns on the sensors. First, the user is stationary during the machine usage. Second, there is a fluctuation in the magnetic field readings as soon as the user interacts with the machine. This fluctuation is due to the distortion from metals and electronic chips installed in these machines forming a peak in the magnetic field readings (detected by the peak detector). Finally, as these machines are usually mounted to walls, the passenger is forced to change her direction to walk away as soon

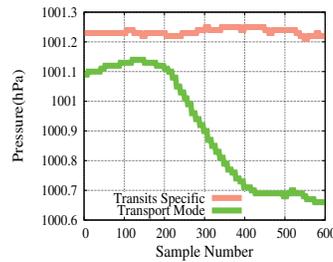


Fig. 3: Comparing barometer readings while changing floor (e.g., elevator) against on the same floor (e.g., entrance gate).

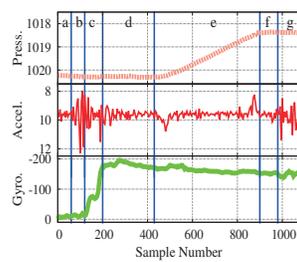


Fig. 4: Elevator usage pattern: (a) waiting for it, (b) walking into it, (c) direction change, (d) stationary, (e) going up, (f) stopping, (g) walking out.

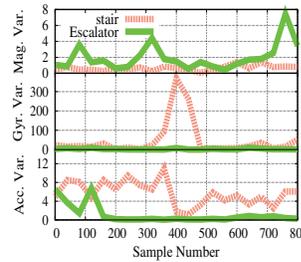


Fig. 5: The sensor patterns that compare climbing up a half-landing stairs against standing on a moving up escalator.

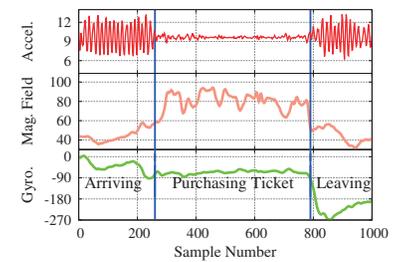


Fig. 6: The acceleration, ambient magnetic field and gyroscope readings while using a coin operated machine.

as she finishes the service. This instantaneous change in the user direction is reflected to a surge in the gyroscope readings when the user starts to resume walking (detected by the surge detector). This unique patterns are leveraged by *LandmarkSense* to separate this type of machines from other landmarks (Figure 2b). Now we will give the detail of how to discriminate the three classes of coin operated machines.

Drink Vending Machine: From the preliminary experiment, we noticed that drink vending machines have a unique loud sound emitted when they are dispensing drinks. This sound is emanated when the drink is pulled down from the machine storage into its outlet. So, we revert to the audio recordings in a bid to separate the drink vending machine from the other two classes of coin operated machines. To exploit the unique drink falling audio signal, in our preliminary experiment we recorded an audio clip during the course of buying a drink activity from the vending machine to analyze it. Figure 7 (d) plots the raw audio signal recorded during this activity in the time domain where the drink drop sound started from the 25th second and lasts to the middle of 26th second. We crop the section of the audio signal comprising the drink falling and background noise (6th second) parts and convert the time domain signals to the frequency domain through 512pt Fast Fourier Transform (FFT) (Figures 7(a), 7(c)). Since all coin operated machines have a similar coin insertion sound, we must ensure that this sound does not share any frequency characteristics with the drink falling sound (used as discriminator). So, we also crop the section of the signal comprising the coin insertion sound (from 12th to 16th seconds) and convert it to the frequency domain as depicted in Figure 7(b).

Referring to Figure 7(c), we observed a clear peak at the 350Hz frequency band in the drink falling audio clip while no peak at 350 Hz frequency band are evident neither in coin insertion nor background noise clips (Figures 7(a), 7(b)). We use an empirical threshold of three standard deviations (i.e., 99.7% confidence level of noise) to detect the drink falling acoustic signal. If the received audio signal strengths in 350 Hz frequency band exceeds the threshold, it means that the signal strength is jumped significantly at this frequency band, the system confirms the detection of the drink vending machine.

Ticket Vending Machine: Similar to drink vending machines, we observed that ticket vending machines emit a unique

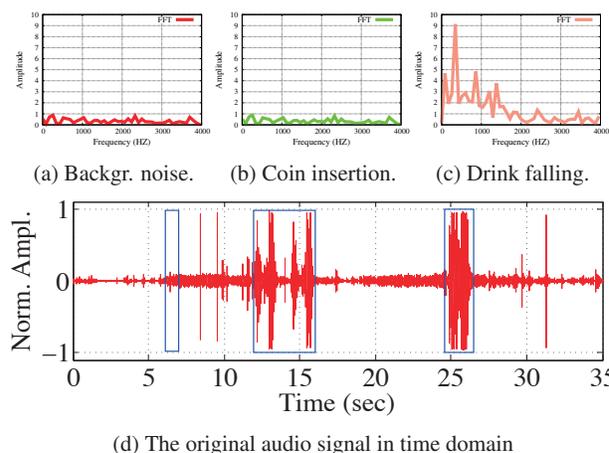


Fig. 7: A sample of audio signal recorded during the usage of a drink vending machine. The raw audio signal in (d) shows three different audio signal bounded by blue boxes corresponding to: the background noise, the coin insertion sound, and the drink falling sound respectively. Figures (a), (b), (c) depict the frequency domain of these three signals respectively.

beep sound many times during the user interaction (e.g. pressing a button, indicating the end of transaction). We envision that this beep signal can be leveraged as a reliable discriminator as it is absent in lockers. We have incorporated the same acoustic detection algorithm used to identify drink vending machines to separate ticket vending machines from lockers. Figure 8(c) shows a raw audio recording collected while a user is using a ticket vending machine. We crop two sections from the original audio signal comprising the background noise and the beep audio signal respectively and convert these signals into the frequency domain as depicted in Figs. 8(a) and (b). We observed a clear peak around the frequency of 3kHz in the beep audio signal whereas no peaks are observed at the frequency of 3kHz in the background noise as shown in Figs. 8(a) and (b). When the ticket vending machine starts beeping, the signal strength in 3kHz frequency band jumps significantly and therefore can be detected using the previous detection algorithm used to identify the drink vending machine.

Lockers: Once vending machines are separated, the remaining samples of coin operated machines are classified as lockers.

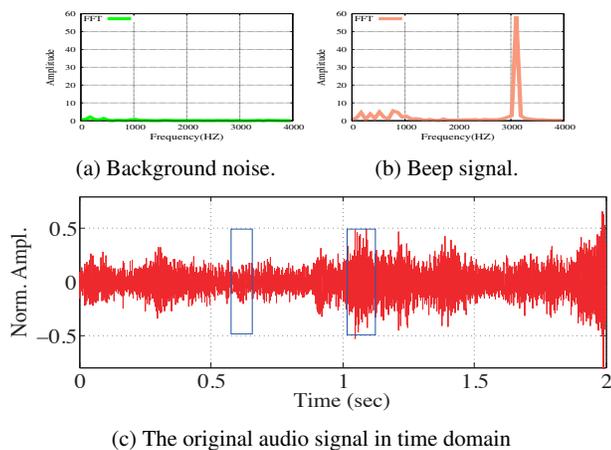


Fig. 8: A sample of audio signal recorded while a user is using a ticket vending machine. The raw audio signal in (c) shows two different audio signals bounded by blue boxes corresponding to the background noise and the beep audio signal respectively. Figures (a) and (b) depict the frequency domain of these two cropped signals respectively.

4.2 Entrance Gate

Railway passengers have to cross an entrance gate in their routes to stations platforms. To cross a gate, there are two methods:

Ticket: Passengers using tickets have to pass by a ticket vending machine beforehand. Thereafter, as a passenger approaches the gate, a noticeable slows down in her walking speed is observed until she pauses in front of the gate to drop the ticket into the machine, then she steps forward to grab it from the machine, and finally she resumes normal walking (Figure 9a). While crossing the gate, there is a distinct peak on the magnetometer readings caused by interference from gate machinery ferromagnetic metals. This unique motion pattern (normal walking, deceleration, accelerating and normal walking) is detected by using the variance of acceleration as depicted in Figure 9a where the two horizontal lines correspond to the thresholds used to separate different motion patterns. The bump on the magnetic field readings is detected by a simple peak detector.

IC Card: Nowadays, IC cards are commonly used for paying transit fees in many areas. This entrance method has two differences from ticket based one. First, passengers using IC cards do not have to pause as the card reader can recognize the card while it is in close proximity in users' hand or in her wallet (acceleration variance still above the stationarity threshold (Figure 9b)). Second, it does not need to be preceded by the usage of a ticket vending machine activity.

5. Performance Evaluation

In this section, we evaluate the accuracy of different landmarks identification, the location accuracy for the discovered landmarks, and finally quantify the the power consumption of *LandmarkSense*. *LandmarkSense* is evaluated through a deployment in major train station having 12 different platforms and covering about 6000 m² area, as well as a major subway station having 5 different platforms, together in the same building complex (Osaka station

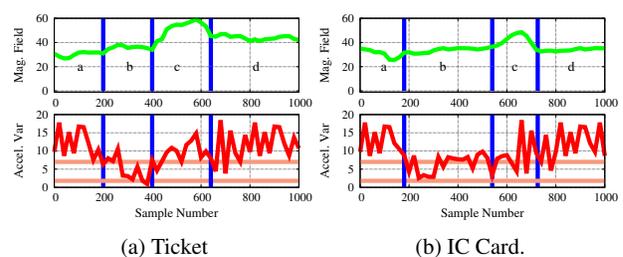


Fig. 9: The sensor pattern of crossing an entrance gate by a ticket and an IC card. Both consist of (a) normal walking, (b) deceleration near the gate, (c) acceleration accompanied by a peak on ambient magnetic field, and (d) normal walking.

city).

5.1 Data Collection Methodology

A group of 9 student volunteers collected the necessary data for evaluation. The participants were assigned specific trajectories starting from the station entrance to different platforms. The trajectories were selected carefully to cover all possible traces that were exhibited by daily passengers while covering all available landmarks at the same time.

During the course of the experiment, participants carried Nexus 5 phones on their hand. We have deployed two Android applications that runs on Android SDK 4.4. The first application is a data collection tool that runs in the background to sample all inertial sensors and the barometer at 50Hz, record audio at sampling rate of 44100Hz. The second application is designed for ground truth collection and runs in the foreground to allow participants to manually timestamp their activities.

5.2 Performance Results

5.2.1 Transport Mode Landmarks Detection

Table 1 shows a confusion matrix for detecting various transport mode landmarks. The matrix shows that most transport mode landmarks are easy to detect due to their unique patterns. This leads to zero false positive and false negative rates for the elevators, half landing stairs, and escalator (when a passenger is standing) cases. Only straight stairs are sometimes misclassified as escalators (when passengers climb them) when their locations are very close to an escalator so they have a magnetic distortion signature similar to escalators. Nevertheless, standing in or climbing up an escalator activities are translated to the same landmark (escalator) and *LandmarkSense* can still achieve a high accuracy with an overall 0.6% false positive and 1% false negative rates.

5.2.2 Stations Specific Landmarks Detection

The confusion matrix in Table 2 shows that vending machines (the coarse-grained category) can be detected with a 100% accuracy using their unique inertial sensors pattern. To classify vending machines to their fine-grained categories (drink and ticket), the acoustic based detection can achieve a good accuracy with an average of 8.6% and 4%, for false positive and negative rates respectively. The table, also, shows that crossing entrance gates by IC cards detection is challenging as some passengers do not slow down sufficiently (misclassified as walking) while others walk very slowly (misclassified as entrance by ticket). However, since

Table 1: Confusion Matrix for classifying different **transport mode** landmarks.

	Elevator	Stairs (Straight)	Stairs (Half Landing)	Escalator (Standing)	Escalator (Climbing)	Escalator (Overall)	FP	FN	Σ
Elevator	41	0	0	0	0	0	0%	0%	41
Stairs (Straight)	0	47	0	0	2	2	0%	4%	49
Stairs (Half Landing)	0	0	16	0	0	0	0%	0%	16
Escalator (Standing)	0	0	0	46	0	-	-	-	46
Escalator (Climbing)	0	0	0	0	37	-	-	-	37
Escalator (Overall)	0	0	0	-	-	83	2.4%	0%	83
Total							0.6%	1%	189

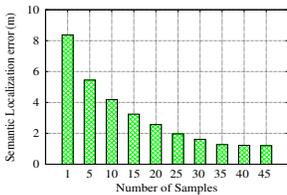


Fig. 10: Effect of the number of samples on the accuracy of semantics location estimation.

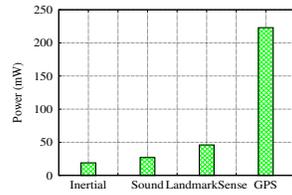


Fig. 11: Energy footprint of *LandmarkSense*.

all gates have IC card readers and ticket slots integrated into the same machine, the two entrance methods (IC card or ticket) are aggregated into one landmark (entrance gate) that can be identified with 4.7% and 7% false positive and false negative rates respectively.

Finally, *LandmarkSense* can consistently detect all types of landmarks accurately with at most 9.7% false positive rate and 7.4% false negative rate.

5.2.3 Discovered Landmarks Location Accuracy

In this subsection, we study how much data is enough for *LandmarkSense* to estimate landmarks' locations accurately as in crowdsensing based systems the accumulation of more samples will enhance the system performance. Figure 10 quantifies the effect of the number of crowd-sensed samples on the accuracy of location estimation. The figure shows that even if some landmarks have some false positive samples, the system can achieve a good accuracy in estimating their locations. This stems from the fact that independent correct samples of the same landmark are in adjacent locations and tend to cluster while erroneous samples are widely scattered in the location space and cannot form a cluster. So, the errors in landmarks location estimation are mainly due to the PDR error. Nevertheless, it is evident from the figure that this error will drop quickly as the number of crowd-sensed samples increased. *LandmarkSense* can consistently achieve the accuracy of 2.5 m using as few as 20 samples and this error is decreased to 1.6m using 35 samples for all discovered landmarks.

5.2.4 Power Consumption

For energy efficiency, *LandmarkSense* leverages low energy inertial sensors to recognize the passenger activities as well as to estimate the passenger's location. In addition, the sound sensor which has higher energy footprint is activated only during time of activities (activities durations are short with 31second in the average excluding the restroom). Thus, *LandmarkSense* has an efficient power consumption profile as shown in Figure 11. To illustrate the power efficiency of *LandmarkSense*, we run an application that samples the GPS every second to show the contrast

in power consumption (even if GPS is neither available in all locations in transits nor it is able to detect landmarks). The power is calculated using the PowerTutor profiler [9] and the android APIs using the HTC Nexus One cell phone. Note that since inertial sensors are used during the normal phone operation, to detect the phone orientation change or estimate the user location for any indoor LBS, *LandmarkSense* practically consumes little extra sensing power in addition to the standard phone operation which results from the intermittent activation of the sound sensor.

6. Related Work

6.1 Mobile Phone Localization

As GPS signal is not available in railway stations, an indoor localization technique is needed to estimate the passenger's location. The most ubiquitous indoor localization techniques are either WiFi-based or dead-reckoning based. WiFi-based techniques, e.g. Horus [10], require calibration to create a prior wireless map for the building. However, the calibration process is time consuming, tedious, and requires periodic updates. Dead-reckoning based localization techniques, e.g. [4], [7], leverage the inertial sensors on mobile phones to dead-reckon the user starting from a reference point [4]. However, dead-reckoning error quickly accumulates leading to complete deviation from the actual path. Therefore, many techniques have been used to reset the dead-reckoning error including snapping to environment anchor points, such as elevators and stairs [7].

LandmarkSense employs the basic concept of Unloc [7] as it provides accurate, energy-efficient localization, and does require neither infrastructure support. However, *LandmarkSense* discovers richer set of transits' landmarks to reset the accumulated localization error.

6.2 Human Activity Recognition

Activity recognition literature has demonstrated the ability to recognize user behavior using either worn-body sensors or smartphone equipped sensors. Accelerometer data was used to detect when a user is walking, standing, running, climbing up the stairs, vacuuming and brushing teeth [11], [12]. Moreover, accelerometer data is used to detect more complex human activities like climbing, biking, driving, lying, cleaning kitchen, cooking, sweeping, washing hands, and medication [13]. Ref. [14] proposes a similar approach to estimate a person's low-level activities and spatial context using data collected by a small wearable sensor device. Furthermore, smartphone sensors have also been leveraged to detect the phone's context. For instance, ambient sensors like temperature, humidity, pressure and light have been used to label user's location directly as being in kitchen, bedroom,

Table 2: Confusion Matrix for classifying station specific landmarks.

	Drink Ven.	Ticket Ven.	Locker	Stationary	Walking	Ent. Gate (Ticket)	Ent. gate (IC)	Ent. gate (Over.)	FP	FN	Σ
Drink Ven.	36	0	0	0	0	0	0	0	8.3%	0%	36
Ticket Ven.	3	69	3	0	0	0	0	0	5.3%	8%	75
Locker	0	4	28	0	0	0	0	0	18.7%	12.5%	32
Stationary	0	0	3	34	0	0	0	0	8.1%	8.1%	37
Walking	0	0	0	0	61	2	4	6	13.4%	8.9%	67
Ent. Gate (Ticket)	0	0	0	0	4	60	2	-	-%	-%	66
Ent. gate (IC)	0	0	0	0	5	3	53	-	-%	-%	61
Ent. gate (Over.)	-	-	-	-	9	-	-	118	4.7%	7.0%	127
Total									9.7%	7.4%	368

bathroom and living room [15]. Moreover, the AmbientSense system [16] can recognize 23 different contexts (e.g. coffee machine, raining, restaurant, dishwasher, street, toilet flush, etc) by analyzing ambient sounds sampled from smartphone. In addition, the RoomSense system in [17] uses active sound probing to classify the type of room (e.g. corridor, kitchen, lecture room) where the user is currently located.

LandmarkSense recognizes a set of passengers’ activities that are mined to discover a richer set of stations landmarks.

6.3 Floorplan Construction

Recently, a number of systems have been proposed that employ pedestrian motion traces to automatically construct indoor floorplans. For instance, CrowdInside [1] processes inertial motion traces using computational geometry techniques to extract the overall floorplan shape as well as corridors and room boundaries. It also identifies a variety of points of interest in the environment such as elevators and stairs. However, their landmark detection method neither targets stations specific landmarks (e.g., entrance gate, etc) nor it provides fine-grained classes of elevation change landmarks (e.g., stair types and elevator types). In addition, their elevator detection algorithm leverages only the motion pattern (Accelerate-Constant-Decelerate) which may coincide with normal human walking patterns. This cannot happen in our method as normal walking traces are separated beforehand using the landmark type detection module. Finally, different passengers behaviors (e.g., climbing up or standing in escalators) makes the low acceleration variance, used in their method, is not a reliable discriminator between climbing stairs and escalators. Jigsaw [3] uses a computer vision approach to extract the position, size, and orientation of landmark objects from images taken by users. It then combines user mobility traces and locations where images are taken to produce the hallway connectivity and the room size. The system proposed in [2] leverages Wi-Fi fingerprints and user motion information to determine which rooms are adjacent in the building and estimating their sizes. It then orders them along each hallway and adjusts the room sizes to optimize the overall floorplan layout. *Nevertheless, all previous systems did not attach any semantic information to the floorplan layout.*

LandmarkSense assumes in its operation the availability of an unlabeled station floorplan, which can be automatically constructed using these approaches. It then enriches the input floorplan with different landmarks based on data collected from users’ phones.

7. CONCLUSION

We presented *LandmarkSense*: a system for automatically enriching railway indoor maps via a crowdsensing approach based on standard cell phones. We implemented *LandmarkSense* using commodity mobile phones running the Android operating system and evaluated it in major train and subway stations. Our results show that we can detect stations landmarks accurately with at most 9.7% false positive and 7.4% false negative rates for all types of landmarks.

References

- [1] Alzantot, M. and Youssef, M.: Crowdsinside: automatic construction of indoor floorplans, *SIGSPATIAL*, ACM (2012).
- [2] Jiang, Y., Xiang, Y., Pan, X., Li, K., Lv, Q., Dick, R., Shang, L. and Hannigan, M.: Hallway based automatic indoor floorplan construction using room fingerprints, *UbiComp*, ACM (2013).
- [3] Gao, R., Zhao, M., Ye, T., Ye, F., Wang, Y., Bian, K., Wang, T. and Li, X.: Jigsaw: Indoor floor plan reconstruction via mobile crowdsensing, *MobiCom*, ACM (2014).
- [4] Jin, Y., Toh, H., Soh, W. and Wong, W.: A robust dead-reckoning pedestrian tracking system with low cost sensors, *PerCom*, IEEE (2011).
- [5] Mohan, P., Padmanabhan, V. and Ramjee, R.: Nericell: rich monitoring of road and traffic conditions using mobile smartphones, *SenSys*, ACM (2008).
- [6] Cleveland, W. and Devlin, S.: Locally weighted regression: an approach to regression analysis by local fitting, *Journal of the American Statistical Association*, Vol. 83, No. 403 (1988).
- [7] Wang, H., Sen, S., Elgohary, A., Farid, M., Youssef, M. and Choudhury, R.: No need to war-drive: Unsupervised indoor localization, *MobiSys*, ACM (2012).
- [8] Ester, M. and et al.: A density-based algorithm for discovering clusters in large spatial databases with noise., *KDD* (1996).
- [9] Gordon, M., Zhang, L., Tiwana, B., Dick, R., Mao, Z. and Yang, L.: PowerTutor: a power monitor for android-based mobile platforms (2013).
- [10] Youssef, M. and Agrawala, A.: The Horus WLAN location determination system, *MobiSys*, ACM (2005).
- [11] Ravi, N., Dandekar, N., Mysore, P. and Littman, M.: Activity recognition from accelerometer data, *AAAI* (2005).
- [12] Kwapisz, J., Weiss, G. and Moore, S.: Activity recognition using cell phone accelerometers, *ACM SigKDD Explorations Newsletter*, Vol. 12, No. 2 (2011).
- [13] Dernbach, S., Das, B., Krishnan, N., Thomas, B. and Cook, D.: Simple and complex activity recognition through smart phones, *Intelligent Environment*, IEEE (2012).
- [14] Subramanya, A., Raj, A., Bilmes, J. and Fox, D.: Recognizing activities and spatial context using wearable sensors, *arXiv preprint arXiv:1206.6869* (2012).
- [15] Mazilu, S. and Troster, G.: A study on using ambient sensors from smartphones for indoor location detection, *WPNC*, IEEE (2015).
- [16] Rossi, M., Feese, S., Amft, O., Braune, N., Martis, S. and Troster, G.: AmbientSense: A real-time ambient sound recognition system for smartphones, *PERCOM Workshops*, IEEE (2013).
- [17] Rossi, M., Seiter, J., Amft, O., Buchmeier, S. and Tröster, G.: RoomSense: an indoor positioning system for smartphones using active sound probing, *Proceed. of the 4th Augmented Human Inter. Conf.*, ACM (2013).