Tracking Motion Context of Railway Passengers by Fusion of Low-Power Sensors in Mobile Devices

Takamasa Higuchi Hirozumi Yamaguchi Teruo Higashino
Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka, Suita, Osaka 565-0871, Japan
{t-higuti, h-yamagu, higashino}@ist.osaka-u.ac.jp

ABSTRACT

In this paper we develop *StationSense*, a novel mobile sensing solution for precisely tracking temporal stop-and-go patterns of railway passengers. While such motion context serves as a promising enabler of various traveler support systems, we found through experiments in a major railway network in Japan that existing accelerometer-based passenger tracking systems can poorly work in modern trains, where jolts during motion have been dramatically reduced. Towards robust motion tracking, StationSense harnesses characteristic features in ambient magnetic fields in trains to find candidates of stationary periods, and subsequently filters out false positive detections by a tailored acceleration fusion mechanism. Then it finds optimal boundaries between adjacent moving/stationary periods, employing unique signatures in accelerometer readings. Through field experiments around 16 railway lines, we show that StationSense can identify periods of train stops with accuracy of 81%, which is almost 2 times higher than the existing accelerometer-based solutions.

Author Keywords

Train stop detection; mobile sensing; multi-sensor fusion

ACM Classification Keywords

I.5.4 Applications: Signal processing

INTRODUCTION

Railways have been an essential transportation infrastructure that supports our daily lives and economic activities, keeping our society *moving*. Especially in urban areas, where population is significantly concentrating, trains would be usually the best way for traveling to one's destination, since other options like buses, cars, and taxis are often affected by heavy traffic jams. According to a survey [11], 79% of commuters in the Tokyo metropolitan area use trains, while 48% of travels by local residents are primarily composed of train rides. This in turn means that improved travel support for such railway passengers (*e.g.*, advanced transfer navigation) has significant potential to optimize human mobility at the city scale.

ISWC '15, September 7–11, 2015, Osaka, Japan.

http://dx.doi.org/10.1145/2802083.2808387

The powerful sensing capabilities of commercial mobile and wearable devices, combined with the recent advances in mobile sensing technology, have been providing promising solutions in this direction, enabling a variety of people-centric mobile applications for railway passengers. One of the most fundamental passenger context is motion state (*i.e.*, moving or stationary) of vehicles. Despite its simplicity, a time series of such binary motion context can help a wide range of mobile-sensing-based passenger support systems:

- **Continuous tracking of railway passengers**. Given the initial location of a passenger by GPS or WiFi positioning systems [2], or by manual input through a mobile app, the phone can continuously track location of the user thereafter by counting the number of stops at stations. Thus it allows people to use navigation apps even under the lack of positioning infrastructures (*e.g.*, in a subway network) [9].

- **Adaptive transfer navigation**. Assuming that some passengers upload their locations that are estimated by such binary motion tracking, it is possible to track locations of trains around the whole city. Such timely and fine-grained operation information can be utilized for advanced transfer navigation, which considers delays of transit vehicles to minimize the passenger's waiting time at stations [10].

- **Efficient crowd control**. When operation of trains is suspended due to accidents or disaster, stations along the line usually become overcrowded, causing serious risks of crowd accidents. The cooperative transit tracking mechanism above also enables early detection of such significant delays, allowing security bureaus to ask potential passengers to wait at their workplace until operation is resumed, or to consider using alternative routes, much earlier than official announcements by railway operators.

A common mobile sensing approach to train stop detection is to employ phone-embedded accelerometers [9, 10]. The idea underlying the existing algorithms is that passenger cars continuously jolt during the train's motion, making large variation in accelerometer readings. On the other hand, technological effort towards better passenger experience has dramatically reduced such jolts. Consequently, as we will show through preliminary experiments, the acceleration noise can keep low level even during vehicle motion, causing considerable errors in accelerometer-based stop detection.

In this paper we design and implement *StationSense*, a novel system for precisely tracking motion context of trains, which

smartphone users ride. To this end, StationSense effectively combines measurements from accelerometers and magnetometers in mobile devices by a tailored data fusion algorithm. Analyzing real sensor measurements, we found that activity of motors and electrical power inverters causes a characteristic feature in an ambient magnetic field in passenger cars during their motion. StationSense harnesses this feature to find candidates of stationary periods, and subsequently filters out false positive detections that occasionally happen due to inertia driving of trains by the acceleration fusion mechanism. Finally it determines optimal boundaries between adjacent moving/stationary periods based on characteristic acceleration signatures, which are typically observed exactly when train stops and restarts moving. Through field experiments, we show that StationSense can identify the stationary periods of trains with accuracy of 81%, which is 80-93% higher than the existing accelerometer-based solutions.

In short, our contributions can be summarized as follows:

1. **A comparative study on existing train stop detection algorithms**. We analyze performance of the existing accelerometer-based solutions through preliminary experiments, clarifying challenges in robust passenger tracking in modern railway systems.
2. **Development of a robust passenger tracking system**. We design a novel train stop detection algorithm to cope with the challenges above. To fully optimize the system based on domain knowledge and thorough analysis of real sensor measurements, we build a task-specific tailored algorithm rather than applying general machine learning algorithms.
3. **Field experiments in real railway networks**. We conducted field experiments in 16 railway lines to analyze performance of StationSense under various combinations of device models, ways of carrying devices, and types of train vehicles, showing its effectiveness in practical situations.

ACCELEROMETER-BASED PASSENGER TRACKING SYSTEMS: A COMPARATIVE STUDY

In this section we briefly outline two existing accelerometer-based solutions, and investigate their performance through a preliminary experiment in a major subway network in Japan.

Existing Solutions for Train Stop Detection

To detect stops at stations, SubwayPS [9] employs a hard threshold on magnitude of acceleration. It first filters out gravity components from the raw sensor readings, and then calculates magnitude a_l of the resulting linear acceleration as $a_l = (a_x^2 + a_y^2 + a_z^2)^{1/2}$, where a_x, a_y, and a_z are filtered acceleration components along three axes of the device's local coordinate system. It concludes that the vehicle is in motion if average acceleration magnitude over a 2-second window exceeds a pre-defined threshold th. Otherwise the vehicle is considered to be stationary. To mitigate impact of acceleration noise, it changes the estimated state only if the acceleration feature keeps above (or below) th for more than pre-configured delay periods of 5–7 seconds.

Thiagarajan et al. [10] propose a probabilistic approach to motion detection. Based on magnitude a of raw acceleration samples, they derive probability that a train is moving as:

$$p(mov|a) = \frac{p(a|mov)p(mov)}{p(a|mov) + p(a|stop)} \quad (1)$$

where $p(a|mov)$ and $p(a|stop)$ are conditional probability distributions of the acceleration magnitudes while vehicles are moving and stationary, respectively. They model the empirical distributions $p(a|mov)$ and $p(a|stop)$ as Laplace distributions $f(a|\mu, b) = (1/2b)\exp\{-|a-\mu|/b\}$, where μ is a median of acceleration magnitude a_i in a training dataset, and $b = \sum_i |a_i - \mu|$. $p(mov)$ is a priori probability that the train is moving, and is set to 0.5. The probability of train motion is subsequently averaged over a sliding window of 30 seconds. Assuming that each peak and valley in the sequence of smoothed motion probability corresponds to a single moving period and a stationary period, respectively, they consider a boundary of different motion states exists between each pair of adjacent peaks and valleys. While the authors do not explicitly describe how they search for an optimal boundary, we implement this part by finding a point in time between each peak-valley pair that maximizes difference in average acceleration magnitudes before and after the candidate of boundary.

Preliminary Experiments in an Urban Railway Network

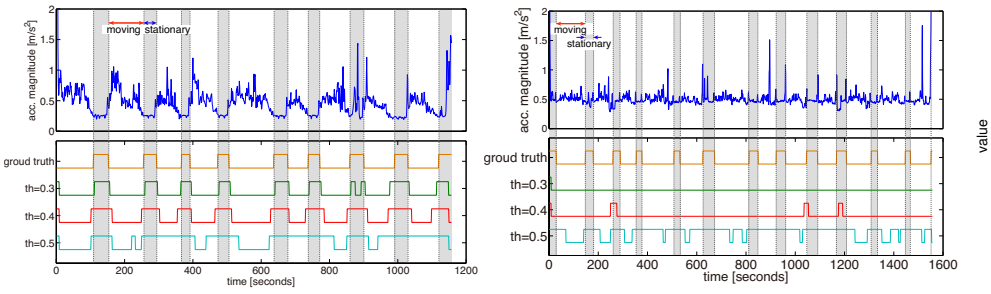
We analyzed performance of the existing solutions above through preliminary experiments using Android smartphones (*i.e.*, Galaxy Nexus). The experiments were conducted in a municipal subway network in Osaka, which consists of 9 electric train lines. We implemented a logger application for the Android OS, which records accelerometer and magnetometer readings at sampling frequency of 60Hz. In the experiment, each participant carries two logger phones: one in a hand and another in a trouser pocket. We also asked the participants to record actual departure/arrival time by tapping buttons on a dedicated annotation app that runs on another smartphone. Through the experiments, we collected annotated sensor data for 18 separate train rides. The resulting dataset includes more than 120 stops at stations and total travel time of more than 8 hours, covering all of the 9 lines.

Performance Metrics

For performance metrics, we consider precision, recall, and f-measure of *train stop detection*: We estimate motion state (*i.e.*, *moving* or *stationary*) of a train every second to obtain a time series of estimated binary state. We define *True Positive* (*TP*) by the total time in seconds when both estimated and ground truth motion state is *stationary*. In the same manner, we define *False Positive* (*FP*) by the total time when the estimated state is *stationary* while the actual state is *moving*, and *False Negative* (*FN*) by the time when *stationary* state is wrongly classified as *moving*. Based on the above definitions, we define the *precision* by $TP/(TP + FP)$, the *recall* by $TP/(TP + FN)$, and the *f-measure* by $(2 \times precision \cdot recall)/(precision + recall)$.

Performance Analysis of Existing Solutions

Figure 1 shows stop detection performance of SubwayPS [9] for train rides on (a) Nanko Port Town Line and (b) Midosuji Line. The upper plot shows a time series of acceleration



(a) (b)

Figure 1. Train stop detection performance by SubwayPS: (a) Nanko Port Town Line and (b) Midosuji Line.

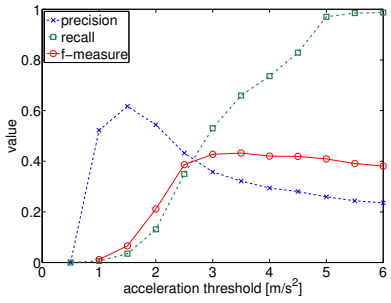
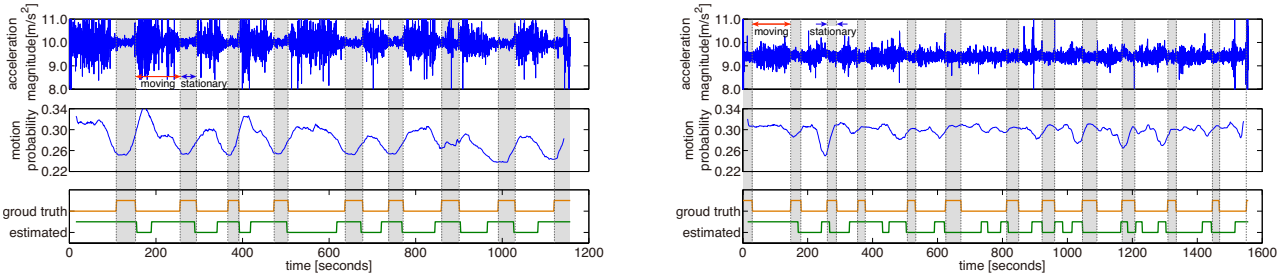


Figure 2. Parameter analysis



(a) (b)

Figure 3. Train stop detection performance by Thiagarajan et al.: (a) Nanko Port Town Line and (b) Midosuji Line.

magnitude, where gravity components are filtered out by the linear accelerometer API of the Android OS. The lower plot shows actual motion state and estimated state calculated by SubwayPS with acceleration thresholds of 0.3, 0.4, and 0.5 m/s^2, where the *stationary* periods are indicated by higher level, while lower level indicates *moving* periods. For the case (a), acceleration magnitude clearly increases over almost the whole train motion because of continuous jolts of the passenger car. By setting th to 0.3 m/s^2, the system can accurately identify the periods of time when the train is stationary. The results for the case (b), however, suggest that the characteristics of linear acceleration can significantly vary depending on train lines. While small acceleration spikes are observed intermittently, similar spikes are also seen in the stationary periods and can hardly be strong evidence of train motion. If we set the acceleration threshold to 0.3 m/s^2, which is the optimal configuration for the case (a), estimated motion state sticks to *moving* for almost the whole trip. Thus performance of SubwayPS is considerably sensitive to the configuration of th. Figure 2 shows stop detection performance by SubwayPS for all train rides in the dataset, when the acceleration threshold is varied between 0.05 and 0.60 m/s^2. Although larger recall can be achieved by pushing up the threshold, it in turn causes increase of false positive stop detections. Consequently, f-measure accuracy cannot be higher than 43% even if the parameter configuration is optimized.

Figure 3 shows stop detection performance of the algorithm in [10] for the same two train rides as in Figure 1. The three plots in the figure show a time series of acceleration magnitude, motion probability derived by an empirical acceleration model, and estimated motion state, respectively. For this experiment, we applied leave-one-out cross validation, in which each train ride in the dataset is sequentially selected as test data, while all the remaining data are used for training the acceleration model. The results show that characteristics of acceleration magnitude is again significantly different between the two cases: while each moving/stationary period in the case (a) contains a single positive/negative peak of the smoothed motion probability, it does not hold in the case (b), causing a considerable number of false positive stop detections. In addition, the smaller variations in the acceleration feature makes it difficult to accurately identify the time of state transition, resulting in large errors around boundaries of adjacent moving/stationary periods.

STATIONSENSE PASSENGER TRACKING SYSTEM

In this section we design the StationSense system based on thorough analysis of the real sensor measurements.

Overview

Figure 4 outlines steps taken by the StationSense system. It first analyzes magnetometer readings to calculate a time series of probability that a train is stationary (*i.e.*, *stop probability*). As we will show in the following section, motors and electrical inverters in vehicles emit large magnetic noise during acceleration periods of the train. Similar magnetic signals are also observed in deceleration periods, since railway systems today usually employ dynamic and/or regenerative braking, both of which slow down the vehicle by converting kinetic energy of wheels into electricity using the motors as power generators: the former just convert the recovered energy into heat using electrical resistors, while the latter feed it back to the power supply. In addition to acceleration and deceleration, there can be inertia driving periods in between, during which both motors and inverters are not active while the train is still in motion. Since electrical and metal materials along railroads form complex spatial magnetic patterns, the

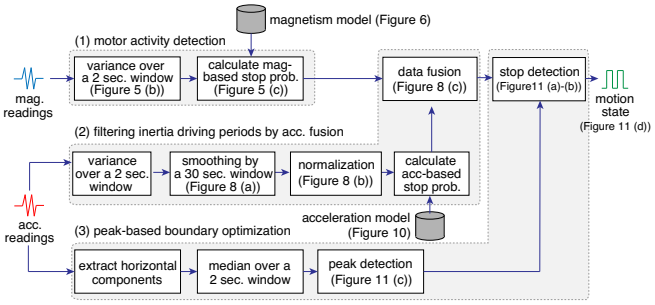


Figure 4. Outline of StationSense passenger motion tracking system

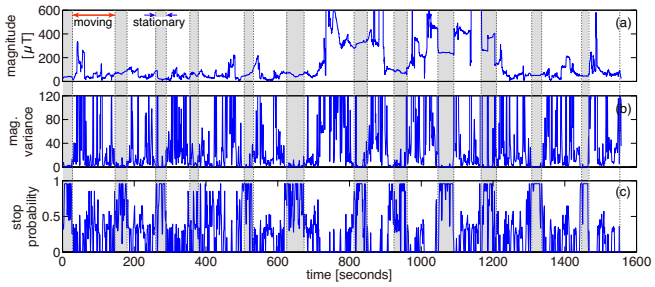


Figure 5. Stop probability based on an ambient magnetic field: (a) magnitude, (b) variance in magnitude, and (c) stop probability.

ambient magnetic field in a passenger car also exhibits characteristic feature as far as the train is in motion. However, difference from stationary periods can be relatively smaller than acceleration and deceleration periods, resulting in potential false positive stop detections. In the subsequent phase we filter out such wrong detections by acceleration fusion, and then search for optimal boundaries between adjacent moving/stationary periods by fusion of the stop probability and a unique acceleration signature, which is typically observed when a train stops and restarts moving.

Motor Activity Detection by Magnetic Field Sensing

Figure 5 (a) shows magnitude of the ambient magnetic field in a passenger car, observed during the same train ride as in Figure 1 (b) and Figure 3 (b). Note that we define magnitude of a magnetic field as $(m_x^2 + m_y^2 + m_z^2)^{1/2}$, where m_x, m_y, and m_z are magnetometer readings along x, y, and z axes, respectively. Although the magnitude values significantly vary over time, it does not necessarily have clear correlation with the train's motion state. In contrast, variance of the magnitude values has totally different characteristics depending upon if the train is in motion. Figure 5 (b) shows variance in magnitude of the magnetic field over a 2-second sliding window. We chose the window size based on typical parameter configurations (*i.e.*, 1-2 seconds) in previous literature [3, 7, 9], and set the sliding margin to 1 second. While large variance values are frequently observed in moving periods, it sticks to fairly low level for almost the whole stationary periods. Figure 6 shows cumulative distributions of the variance values for all the dataset collected in our preliminary experiment. In stationary periods, more than 80% of the variance samples are concentrated below 10.0. This is in contrast to moving periods, in which the variance samples are distributed over a much wider range. This suggests that variance in magnitude

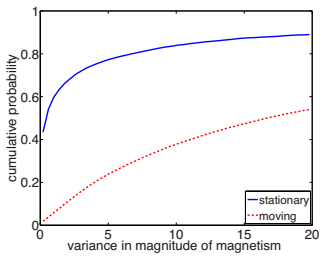


Figure 6. Cumulative distributions of $P(m|stop)$ and $P(m|mov)$

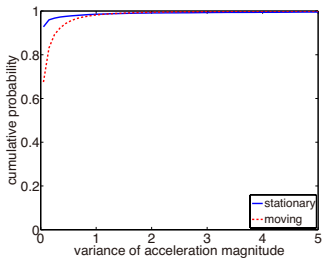


Figure 7. Cumulative distributions of acceleration variance

of the magnetic field serves as a powerful feature for classifying motion state of trains.

Given a variance sample m, stop probability of a train can be calculated based on the magnetism models above as:

$$p(stop|m) \stackrel{\text{def}}{=} \frac{p(m|stop)}{p(m|stop) + p(m|mov)}. \quad (2)$$

We calculate the stop probability based on each variance sample to find a set of stationary periods. Figure 5 (c) shows the resulting stop probability for this train ride. The probability keeps around 1.0 while the train is stationary, which is much higher than those in moving periods. However, there can be temporary increase in stop probability also in some moving periods due to inertia driving of trains. We mitigate the risk of potential false positive stop detections in the following phase.

Filtering Inertia Driving Periods by Acceleration Fusion

The jolts of passenger cars would, at least occasionally, happen throughout a train's motion. The basic idea for our noise filtering mechanism is to harness such long-term temporal patterns in accelerometer readings to remove the undesirable rise in stop probability in the middle of moving periods.

To define an acceleration feature for noise filtering, we first calculate variance of acceleration magnitude over a 2-second sliding window. Figure 7 shows cumulative distribution of the variance values, which is again modeled using the dataset from our preliminary experiment. While the resulting variance values occasionally exhibit large spikes, most of the variance samples are concentrated below 1.0 for both moving and stationary periods. This means that acceleration variation caused by trains' motion usually fall into this range. Thus we first filter out large variance samples greater than 1.0, assuming that they are noise produced by motion of user's body. Since intensity of jolts can be different among individual railway lines, effectiveness of noise filtering could be further improved by train-dependent parameter tuning at the cost of initial calibration effort. In this paper, however, we employ a common threshold to remove the need for such calibration, and conservatively choose the parameter to suppress the risks that jolts of vehicles are filtered out unexpectedly.

Since high frequency components in the sequence of acceleration variance are also usually dominated by noise due to motion of user's body, we subsequently average the filtered variance samples over a long sliding window of 30 seconds. Note that the window size is empirically determined through parameter analysis over our Osaka Municipal Subway dataset,

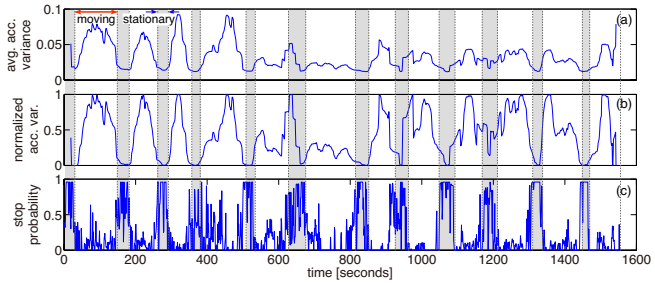


Figure 8. Filtering out inertia driving periods by acceleration fusion: (a) smoothed acceleration variance, (b) normalized acceleration variance, and (c) stop probability after filtering.

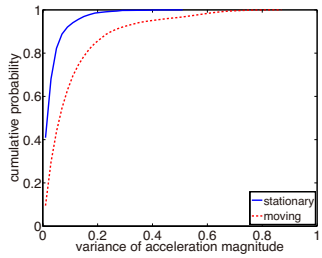


Figure 9. Cumulative distributions of smoothed acceleration variance

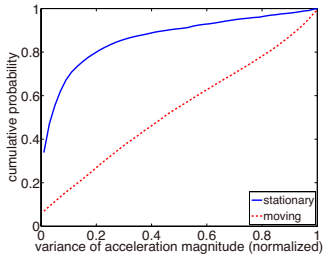


Figure 10. Cumulative distributions of $P(\overline{a}|stop)$ and $P(\overline{a}|mov)$

and we set the sliding margin to 1 second to update the acceleration feature every second. Figure 8 (a) shows the resulting smoothed variance for the ride we have been discussing earlier. While the variance values take a positive peak in a moving period and a negative peak in a stationary period in most cases, the peak values are not stable even during a single ride on the same train, causing ambiguity in the acceleration-based motion classification model. To cope with the problem, we first find peaks in the smoothed acceleration variance by Algorithm 1, and subsequently normalize it into the range between 0 and 1 based on the most recent peak values, as outlined in Algorithm 2. We set the parameter δ for the peak detection to 0.03, which on average achieved reasonable performance in preliminary parameter analysis on our subway dataset. Figure 8 (b) shows smoothed acceleration variance after the normalization process, say $\overline{a}$, while Figures 9 and 10 show cumulative distributions of the acceleration feature before and after normalization, respectively. The variation in the peak values is successfully canceled out after normalization, making distributions of the resulting acceleration feature clearly different according to the train's motion state.

We use this acceleration model for filtering out the rises in stop probability during inertia driving periods. Assuming that the magnetism-based feature m and the acceleration feature $\overline{a}$ are probabilistically independent, we fuse these distributions by the Naïve Bayes model [6]:

$$p(stop|m,\overline{a}) \stackrel{\text{def}}{=} \frac{p(m|stop)p(\overline{a}|stop)}{p(m|stop)p(\overline{a}|stop)+p(m|mov)p(\overline{a}|mov)}. \quad (3)$$

Subsequently we update the stop probability as follows:

$$p(stop) \stackrel{\text{def}}{=} \min(p(stop|m), p(stop|m,\overline{a})). \quad (4)$$

Algorithm 1 Peak detection

```
find_max ← false, mx ← −∞, mn ← ∞
for each smoothed acceleration sample a do
    if a > mx then mx ← a
    if a < mn then mn ← a
    if find_max then
        if a < mx − δ then
            mn ← a, find_max ← false, output mx as a positive peak
    else
        if a > mn + δ then
            mx ← a, find_max ← true, output mn as a negative peak
```

Algorithm 2 Peak-based adaptive normalization

```
Initialize max_var and min_var by pre-defined parameters
for each smoothed acceleration sample a do
    find peaks in smoothed acceleration variance
    if positive peak is detected then max_var ← peak value
    if negative peak is detected then min_var ← peak value
    ā ← min(max{0, (a − min_var)/(max_var − min_var)}, 1)
```

Note that we upper bound the stop probability by the original magnetism-based probability $p(stop|m)$ rather than directly employing $p(stop|m,\overline{a})$ for stop detection, because the normalized acceleration variance can occasionally take relatively low values in some moving periods. In that case, the stop probability during motion can be rather pushed up after acceleration fusion, potentially causing additional false positives. Figure 8 (c) shows the filtered probability for the example ride, where stop probability in moving periods is successfully suppressed, compared to Figure 5 (c).

Stop Detection and Boundary Optimization

Given the sequence of stop probability in Equation (4), StationSense finally identifies a set of stationary periods.

The system first averages the sequence of stop probability over a sliding window of 30 seconds and then detects peaks in the resulting smoothed probability curve. As shown in Figure 11 (a), each stationary period usually contains a single positive peak, for each of which we search for optimal boundaries with its adjacent moving periods.

For boundary optimization, we consider another sliding window of 40 seconds long, centered at each time step. The window size is again empirically determined based on the Osaka Municipal Subway dataset. We then calculate difference in average stop probability over the former and the latter half of the window to find the time when the probability changes the most. Figure 11 (b) shows the resulting differential probability for the same example ride. Negative peaks and positive peaks appear at almost the same time as the beginning and the end of each actual stationary period, respectively.

Analyzing the dataset, we also found that there is usually a spike in magnitude of horizontal acceleration components exactly when trains stop at and depart from stations. StationSense reduces the solution space for optimal boundaries by finding such peaks: It extracts gravity components in acceleration samples by the algorithm in [3], and calculates the components that are orthogonal to the estimated gravity vector. Since the resulting horizontal acceleration vector contains a large amount of high frequency noise, we subsequently

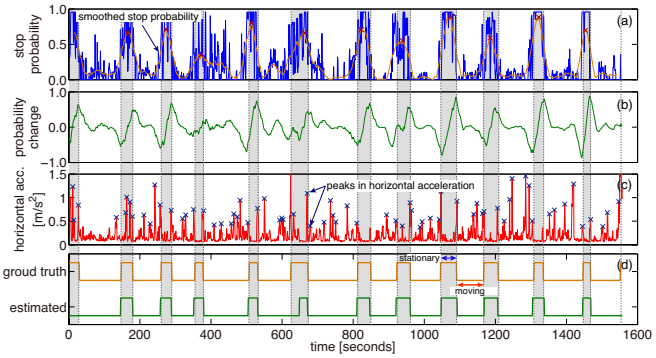


Figure 11. Boundary optimization: (a) stop probability, (b) change in stop probability, (c) horizontal acc., and (d) estimated motion state.

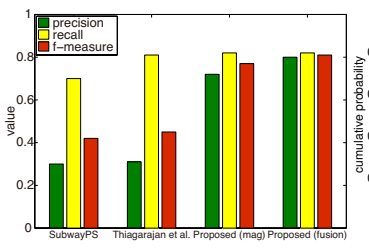


Figure 12. Performance comparison

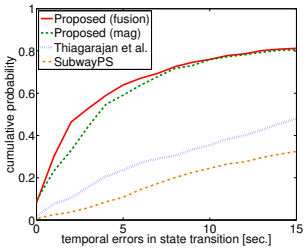


Figure 13. Temporal errors

smooth it by a median filter over a 2-second window. Finally we find positive peaks in the smoothed magnitude of horizontal acceleration, as shown in Figure 11 (c). Then it employs these peaks as candidates of state boundaries.

For finding an optimal former (latter) boundary of a stationary period, StationSense first searches for the time when the differential stop probability is minimized (maximized), and considers it as an initial solution for a boundary. It then calculates a list of points in time over a 30-second window centered at the initial solution, at which horizontal acceleration peaks out. If at least one peaks are detected in the window, it regards the one with the largest change in stop probability as the estimated boundary. Otherwise it adopts the initial solution as the final estimate for the state boundary.

Figure 11 (d) shows a time series of motion state, which is estimated by our StationSense system. In contrast to the accelerometer-based solutions, it precisely identifies boundaries of each stationary period, providing fine-grained passenger motion tracking typically in a few seconds granularity.

EVALUATION

To evaluate tracking performance of StationSense, we conducted additional field experiments to collect a larger dataset under a variety of conditions. This section describes details of the experiment, followed by analysis of the evaluation results.

Field Experiment

We conducted measurement experiments in 16 major railway lines, including 9 lines from the Osaka Municipal Subway, 3 lines from the Hankyu Railway, and 1 line each from other 4 local railway companies around Osaka. For the experiment, we recruited 6 volunteers and asked each participant to carry a smartphone (*i.e.*, Galaxy Nexus and Galaxy Note II from Samsung Electronics, or ZenFone 5 from ASUSTeK) in a hand or in a trouser pocket. In addition to the logger phone, they also carry another smartphone in their hand for annotation with our dedicated mobile application. The participants freely choose to stand or sit on a seat during their rides, while they are asked to keep the same pose (*i.e.*, standing or sitting) over the whole single ride for ease of annotation. Through the experiments above, we collected sensor readings for totally 66 train rides, composed of 386 stops at stations.

Performance Comparison with Existing Solutions

Figure 12 shows precision, recall, and f-measure accuracy of train stop detection by SubwayPS [9], Thiagarajan et al. [10], and our StationSense system (labeled as *Proposed (fusion)*). To examine effectiveness of acceleration fusion and peak-based boundary optimization, we also show performance of a simplified version of StationSense, which employs only magnetometer readings (labeled as *Proposed (mag)*). In the simplified version, we directly use the magnetism-based stop probability in Equation (5) for the subsequent stop detection and boundary optimization, assuming that state boundaries are at points in time when changes in the magnetism-based stop probability is maximized. For SubwayPS, we set the acceleration threshold th to 0.35 m/s^2, which achieved the highest performance in our parameter analysis in Figure 2. Thiagarajan et al. and StationSense are again evaluated by leave-one-out cross validation, where each train ride in the data set is sequentially selected as test data while the acceleration and magnetism models are trained using all the remaining data.

Both SubwayPS and Thiagarajan et al. suffer from frequent false stop detections due to insufficient intensity of jolts during trains' motion. Consequently, they result in low precision of 30% and 31%, respectively, and f-measure accuracy of 42% and 45%. Performance of *Proposed (mag)* shows effectiveness of our magnetism-based stop detection mechanism, achieving precision of 72% and recall of 82%. Thus the robust magnetism-based feature can even serve alone to achieve much higher accuracy than the existing accelerometer-based solutions. The acceleration fusion mechanism of StationSense further enhances precision by 8% by filtering out false stop detections in inertia driving periods, keeping the equivalent recall of 82%.

Figure 13 shows cumulative distributions of errors in estimated time of state transition: for each time a train stops or starts moving, we find the nearest corresponding transition in the sequence of estimated motion state, and define the error by absolute difference between the actual and estimated transition time. The StationSense can bound the errors in state transition time to less than 10 seconds in about 80% of the cases, while the corresponding percentage for SubwayPS and Thiagarajan et al. are around 20-40%.

Impact of User/Device Poses

To examine impact of poses of users and how they carry mobile devices, we separately evaluated stop detection performance under different combinations of user poses (*i.e.*, standing or sitting on a seat) and phone poses (*i.e.*, held in a hand or kept in a trouser pocket). As with other analyses in this

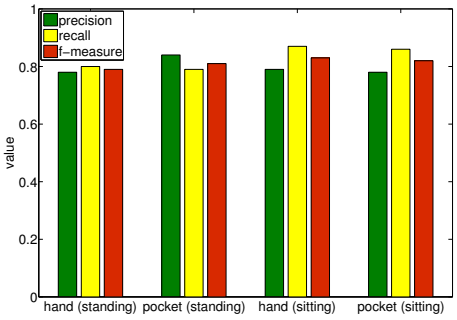


Figure 14. Impact of user/device pose

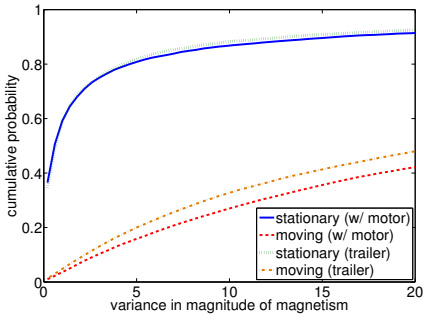


Figure 15. Impact of vehicle types

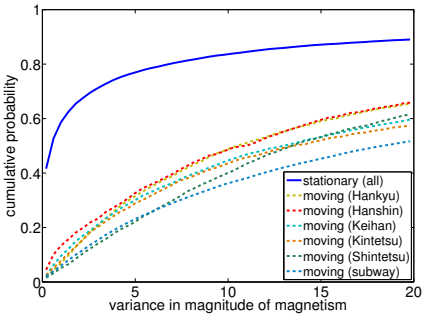


Figure 16. Impact of vehicle models

section, we construct the acceleration/magnetism models by leave-one-out cross validation over the whole data set without any pose-dependent calibration. Figure 14 compares performance under the four different measurement conditions. The results suggest that stop detection accuracy is not significantly dependent on user/device poses, and StationSense can achieve reasonable accuracy in all the cases with f-measure accuracy of about 80%. This robustness mainly comes from the pose-independent feature of the magnetometer readings, which are not largely attenuated by human bodies and relatively stable against motion noise produced by users. While the acceleration fusion mechanism can be potentially affected by such motion noise to some extent, its impact is effectively mitigated by averaging the acceleration variance over a long time window, followed by normalizing the smoothed features.

Impact of Difference in Vehicle Types

Urban trains are usually composed of multiple passenger cars, and some of them may be trailer vehicles which are not equipped with any motors and inverters. To examine the impact of vehicle types on our magnetism-based stop detection mechanism, we conducted the following additional experiment in the Tanimachi Line of the Osaka Municipal Subway. Each volunteer carries a Galaxy Nexus phone in their hand and simultaneously rode different types of passenger cars in the same train. By repeating such experiments 8 times, we formed a dataset, composed of 109 stops at stations. Figure 15 shows cumulative distributions of the magnetism-based feature, which are separately calculated for motor-equipped vehicles and trailer vehicles. The fluctuation of magnetic fields in trailer vehicles tend to be relatively smaller than in those equipped with motors, since magnetic signals emitted from the electric equipment are attenuated with distance. However, there is still significant difference between feature distributions for each state, making the magnetism-based feature a powerful evidence for train stop detection. Consequently, StationSense achieved f-measure accuracy of 84% in both types of vehicles, while precision for motor-equipped vehicles were slightly higher by 1% than in trailer vehicles.

DISCUSSION

Energy Consumption

Energy efficiency is a key design criterion for mobile sensing systems. Hemminki et al. [3] report that battery expenditure for continuous sampling of accelerometer and magnetometer measurements are 73-88% lower than GPS sampling and

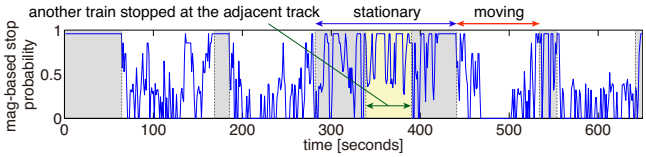


Figure 17. Impact of trains on neighboring tracks

90-96% lower than a phone screen. In addition to the inherent low-power feature of the sensors, there could even be potential for further reduction in energy consumption, if train schedules are available online: once a train's departure is detected, the system could temporarily suspend sensing for a certain time interval, which is sufficiently shorter than standard travel time to the next station. We leave this possible extension for future work.

Delay in Passenger Tracking

StationSense employs sliding window techniques to mitigate impact of noise in sensor measurements. In return for the improved robustness, there is a typical delay of 20 seconds (*i.e.*, a half of the 40 second window for state transition detection) before the system outputs estimated motion state. While we could reduce such delay by adopting shorter windows at the cost of certain accuracy degradation, we believe that delay of a few tens of seconds would be acceptable in many practical passenger support systems discussed in the introduction.

Impact of Difference in Vehicle Models

It would be natural to expect that intensity of magnetic noise during motion can vary depending on vehicle models. Figure 16 shows cumulative distributions of magnetism-based features, separately calculated for each railway operator (*i.e.*, 5 local railway companies and the Osaka Municipal Subway). These operators usually use vehicles made by different sets of manufacturers, while there can be a certain amount of overlap. Although the feature distributions in *moving* state vary among operators, all of them have significant difference from those in *stationary* state, making the magnetism-based feature effective for stop detection.

Magnetic Noise from Trains on Neighboring Tracks

Figure 17 shows a time series of magnetism-based stop probability observed during a ride on the Hanshin Railway. At the middle of the long stationary period around 280-440 seconds, there was another train having a brief stop at the opposite side of the platform. Consequently, stop probability for

the stationary train temporarily dropped due to the magnetic noise, caused by acceleration/deceleration of the neighboring train. We suspect that such magnetic noise from other trains would be a major error-inducing factor, which caused false negative stop detections in our evaluation. A possible way to mitigate such errors would be noise filtering in a temporal domain (*e.g.*, filtering out short-term drops in stop probability). Further analysis on this issue is also an important piece of our future work.

RELATED WORK

A body of research has focused on transportation mode estimation for accurate monitoring of users' mobility behavior [7]. Hemminki et al. [3] employ accelerometers in mobile devices to classify user context into stationary, walk, bus, train, metro, and tram. Analyzing a variety of acceleration features, they find that horizontal acceleration components offer effective features for characterizing vehicular mobility. Sankaran et al. [8] use phone-embedded barometers for transportation mode estimation to remove impact of acceleration noise produced by phone holders. These works are orthogonal to ours, which aims for accurate motion tracking of railway passengers. Note, however, that we could use these systems to automatically start passenger tracking when users get on a train.

Some recent works propose mobile sensing solutions for real-time tracking of transit vehicles. Zhou et al. [12] developed a crowd-sourced bus location system to predict arrival time of buses at each bus stop. For energy efficient vehicle tracking, they employ a time series of received signal strength from cell towers, which is observed by passengers' mobile phones along their bus routes. EasyTracker [1] enables similar bus tracking system by analyzing GPS traces collected from mobile phones installed in each bus. Although these systems can be, in theory, also applied to tracking of *railway* passengers, radio signals from cell towers and GPS satellites may not be always accessible during train rides (*e.g.*, in a subway). Maier et al. [5] develop a system to recognize context of subway passengers by combination of phone-embedded motion sensors and microphones. In return for fine-grained recognition of 17 different activities, it requires continuous sampling and processing of ambient audio, making the system power-hungry. Stockx et al. [9] and Thiagarajan et al. [10] developed accelerometer-based solutions, employing jolts of trains to detect their motion. Although they successfully remove dependence on any positioning infrastructures (*e.g.*, GPS), and at the same time achieve energy efficiency, the potentially small jolts of passenger cars and acceleration noise produced by device holders can seriously harm reliability of the system. Lee et al. [4] suggest possibility of employing magnetometers for train stop detection. Their finding is that intensity of magnetometer readings peaks out soon after the trains' departure. Thus they search for positive peaks in the intensity values and regard the previous negative peak as the departure time. However, they also report that this characteristic may not always hold. The problem can also be seen in our data in Figure 5 (a), where intensity values do not necessarily have clear correlation with the train's motion state. StationSense copes with the problems above by a robust magnetism-based feature and the tailored data fusion algorithm.

CONCLUSION

In this paper we have proposed *StationSense*, a novel mobile sensing solution for precisely tracking motion context of railway passengers. It harnesses characteristic feature in ambient magnetic fields in trains to coarsely identify their motion state, and subsequently applies a tailored acceleration fusion mechanism for accuracy enhancement. Through field experiments, we have shown that StationSense can identify stationary periods at stations with much higher accuracy than existing accelerometer-based passenger tracking systems.

ACKNOWLEDGEMENTS

This work was supported in part by JSPS KAKENHI Grant Number 26220001.

REFERENCES

1. Biagioni, J., Gerlich, T., Merrifield, T., and Eriksson, J. Easytracker: Automatic transit tracking, mapping, and arrival time prediction using smartphones. In *Proc. SenSys* (2011), 68–81.
2. Chintalapudi, K., Padmanabha Iyer, A., and Padmanabhan, V. N. Indoor localization without the pain. In *Pro. MobiCom* (2010), 173–184.
3. Hemminki, S., Nurmi, P., and Tarkoma, S. Accelerometer-based transportation mode detection on smartphones. In *Proc. SenSys* (2013), 13:1–13:14.
4. Lee, G., and Han, D. Subway train stop detection using magnetometer sensing data. In *Proc. IPSN* (2014).
5. Maier, M., and Dorfmeister, F. Fine-grained activity recognition of pedestrians travelling by subway. In *Proc. MobiCASE* (2013), 122–139.
6. Murphy, K. P. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.
7. Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., and Srivastava, M. Using mobile phones to determine transportation modes. *ACM Trans. Sen. Netw. 6*, 2 (2010), 13:1–13:27.
8. Sankaran, K., Zhu, M., Guo, X. F., Ananda, A. L., Chan, M. C., and Peh, L.-S. Using mobile phone barometer for low-power transportation context detection. In *Proc. SenSys* (2014), 191–205.
9. Stockx, T., Hecht, B., and Schöning, J. SubwayPS: Towards smartphone positioning in underground public transportation systems. In *Proc. SIGSPATIAL* (2014), 93–102.
10. Thiagarajan, A., Biagioni, J., Gerlich, T., and Eriksson, J. Cooperative transit tracking using smart-phones. In *Proc. SenSys* (2010), 85–98.
11. Tokyo Urban Area Traffic Plan Meeting. The 5th Tokyo urban area person trip survey. `http://www.tokyo-pt.jp/`.
12. Zhou, P., Zheng, Y., and Li, M. How long to wait?: Predicting bus arrival time with mobile phone based participatory sensing. In *Proc. MobiSys* (2012), 379–392.