Accelerometer-Based Transportation Mode Detection on Smartphones

Samuli Hemminki, Petteri Nurmi, Sasu Tarkoma
Helsinki Insitute for Information Technology HIIT
PO Box 68, Department of Computer Science
FI-00014, University of Helsinki, Finland
firstname.lastname@cs.helsinki.fi

ABSTRACT

We present novel accelerometer-based techniques for accurate and fine-grained detection of transportation modes on smartphones. The primary contributions of our work are an improved algorithm for estimating the gravity component of accelerometer measurements, a novel set of accelerometer features that are able to capture key characteristics of vehicular movement patterns, and a hierarchical decomposition of the detection task. We evaluate our approach using over 150 hours of transportation data, which has been collected from 4 different countries and 16 individuals. Results of the evaluation demonstrate that our approach is able to improve transportation mode detection by over 20% compared to current accelerometer-based systems, while at the same time improving generalization and robustness of the detection. The main performance improvements are obtained for motorised transportation modalities, which currently represent the main challenge for smartphone-based transportation mode detection.

Categories and Subject Descriptors

I.5.2 [**Pattern Recognition**]: Design Methodology: Feature evaluation and selection; I.5.4 [**Pattern Recognition**]: Applications: Signal processing; H.4.m [**Information Systems**]: Information Systems Applications: Miscellaneous

General Terms

Algorithms, Experimentation

Keywords

Mobile Sensing, Activity Recognition, Transportation Mode Detection

1. INTRODUCTION

The increased sensing capabilities of contemporary smartphones combined with their easy programmability, large market penetration rate, and effective distribution channels for third party applications, have resulted in smartphones maturing into an effective tool for unobtrusive monitoring of human behavior [21]. This paper focuses on a specific aspect of human behavior, the transportation behavior of individuals. The capability to capture transportation behavior accurately on smartphones would have a positive impact on many research fields. For example, human mobility tracking would directly benefit from an ability to automatically monitor the transportation behavior of individuals [18, 28]. This in turn would enable improving urban planning [38], monitoring and addressing the spread of diseases and other potential hazards, as well as providing emergency responders information of the fastest route to aid the lost or injured [29]. Localization and positioning algorithms could be improved by constructing more elaborated motion models with the help of information of the user's current transportation mode [25] or by constricting the plausible locations of the user to the route of the detected transportation mode. Persuasive applications could use the transportation behavior monitoring to automatically calculate, for example, CO_2-footprint or level of physical activity [9]. Finally, transportation monitoring could be used as part of user profiling, e.g., for real-time journey planning and guidance systems, or targeted advertising.

While the idea to use smartphones for monitoring transportation behavior itself is not new (see Sec. 2), previous work has primarily focused on elaborate use of the phone's integrated GPS receiver. While GPS-based systems can be very efficient when GPS signals are available, they suffer from some important limitations. First, integrated GPS receivers are well-know to suffer from high power consumption, which means these approaches rapidly deplete the battery of the mobile device, especially when the user is mobile. Second, the GPS receiver's dependency on unobstructed view to satellites presents problems in many common situations of urban transportation, e.g., while the user is moving underground, inside a station, along urban canyons, or while the user is traveling in a vehicle but is unable to stay sufficiently close to a window. Third, current GPS-based solutions provide only modest accuracy when a fine-grained distinction of motorised transportation modes is required. Distinguishing reliably between different motorised transportation modes would provide more detailed information about human transportation behavior, e.g., enabling to au-

tomatically estimate the carbon footprint of individuals or to obtain a detailed understanding of the commuting patterns of urban citizens.

In this paper we present novel accelerometer-based techniques which can be used individually, or in conjunction with other sensors for transportation mode detection on smartphones. We focus on accelerometers as they are well-suited to overcome the above mention limitations. First, accelerometers have very low power consumption, enabling continuous transportation behavior monitoring. Second, accelerometers measure user's movements directly and therefore do not depend on any external signal sources. Third, accelerometers contain highly detailed information about phone movement, enabling fine-grained distinction of different motorised transportation modalities.

A central challenge in accelerometer-based transportation mode detection is to distinguish information pertaining to movement behavior from other factors that affect the accelerometer signals. In particular, gravity, user interactions and other sources of noise can mask the relevant information. As our first technical contribution, we describe a novel gravity estimation technique for accelerometer measurements. Our gravity estimation technique provides more accurate and robust gravity component estimation during motorised transportation, which in turn translates to more accurate horizontal accelerometer representation. The horizontal acceleration is a key factor for identifying motorised modalities, as the acceleration/deceleration periods are typically similar within the same modality, but also distinct from other modalities. The real-world analogy is that different types of vehicles can be identified from their acceleration and breaking periods. As our second technical contribution, we introduce a new set of accelerometer features, so-called *peak features*, which characterize acceleration and deceleration patterns during motorized modalities. These features are a key enabler for improving the detection performance of transportation mode detection approaches. Moreover, these features pave way for new types of sensing applications that analyze vehicular movement patterns, e.g., driving style or fuel consumption estimation.

We evaluate our approach using over 150 hours of transportation data collected from 16 individuals and 4 different countries. The results of our evaluation demonstrate that our approach is able to improve detection accuracy by over 20% compared to current accelerometer-based solutions, and even exceed the accuracy of the current state-of-art hybrid GPS and accelerometer system by over 10%. Moreover, our proposed techniques improve the robustness of the detection and generalize well across users and geographical locations.

2. RELATED WORK

Transportation mode detection can be considered a special subfied of activity recognition, a widely studied field within the wearable and ubiquitous computing communities [17]. The first approaches to transportation mode detection relied on custom sensing platforms [1, 5], whereas recent work has predominantly considered smartphones as the sensing platform. In the following we focus exclusively on previous work on smartphone-based transportation mode detection. We refer to [26] for information about approaches that rely on custom sensing platforms.

Transportation mode detection can be subdivided into two main subtasks: (i) determining whether the user is moving; and, in case movement is detected, (ii) what kind of means the person is using for moving around. The former task, stationarity detection, has been widely explored in different domains. For example, the LOCADIO positioning system classifies the user as mobile or stationary based on changes in the WiFi signal environment [16]. Muthukrishnan et al. [24] detect mobility by analyzing spectral characteristics of the WiFi signal environment. Kim et al. [14] use variance of accelerometer values to detect when a person is staying within one place. Kjaergaard et al. [15] combine accelerometer variance with a threshold on GPS speed to separate motorised transportation modalities from stationary behavior.

The latter task, locomotion detection, has been increasingly explored on smartphones. Typical locomotion types include different pedestrian modalities (e.g., walking, running or moving in stairs), non-motorised transportation (e.g., bicycling, roller skating) and motorised transportation (e.g., bus, train or car). In terms of sensors, the accelerometer is the most widely used sensor for detecting locomotion. A number of early systems used the embedded accelerometer for detecting different pedestrian and non-motorised modalities, such as walking and running [20, 12], ascending or descending stairs [4] or cycling [2]. Wang et al. [31] compare features extracted from the L_2 norm with features extracted from horizontal and vertical representations. From each representation, an extensive set of features is extracted over 8 second non-overlapping windows, and classification is performed using a decision tree. In contrast to our work, the authors obtain the best results using only the L_2 norm representation, which is due to inaccurate estimation of the gravity component and insufficiently expressive feature space. While previous accelerometer-based systems have been effective at detecting pedestrian and non-motorised transportation modalities, achieving typically over 90% accuracies, their performance has been significantly lower for stationary and motorised transportation modalities [26, 31]. Our work improves on these approaches by demonstrating that capturing features from the acceleration/deceleration periods from vehicular motion can be used for accurate and fine-grained detection of motorised transportation modalities.

Instead of relying on the accelerometer, Zheng et al. [36, 37] detect transportation modalities using features extracted from GPS measurements. In addition to speed and location information, the authors consider features that characterize changes in movement direction, velocity and acceleration. Together with information about street segments, the authors reach an average accuracy of 76% in classifying between stationarity, walking, biking, driving and traveling by bus. Recent work has focused on decreasing energy consumption by requiring only sparse GPS data [3], introducing more effective graph-based postprocessing techniques [36], and improving the detection accuracy by fusing in external information on the real-time location of the transportation vehicles [30]. Reddy et al. [26] combine GPS and accelerometer to recognize between stationary, walking, running, biking and motorised transportation, achieving high, over 90% accuracies. Classification is performed with a hybrid classifier consisting of a decision tree and a first order discrete HMM classifier. However, in contrast to our work, Reddy et al. make no distinction between different motorised modalities and mainly rely on GPS speed for detecting motorised transportation. The main drawbacks of all these approaches are that GPS receiver has high power consumption, requires

inconsistent time for obtaining satellite lock, and is unavailable or unreliable when view to satellites is obstructed, e.g., when the user is underground, inside a station, moving in urban canyons or is insufficiently close to a window in a transportation vehicle.

An alternative to GPS is to estimate movement by monitoring changes in the user's signal environment. Sohn et al. [27] use changes in the GSM signal environment for coarse-grained detection of transportation modalities. Mun et al. [23] combine GSM and WiFi for detecting between dwelling, walking and driving, reaching accuracies in the range of $80 - 90\%$. While energy-efficient compared to GPS, these techniques are susceptible to varying WiFi access point density and GSM cell sizes between different locations. Consequently, these techniques are unreliable outside urban areas and require careful calibration, and thus, struggle to generalize to new environments.

3. TRANSPORTATION MODE DETECTION

We have developed a novel solution for transportation mode detection that can provide robust, accurate and fine-grained detection of transportation modalities despite relying solely on the embedded accelerometer of the smartphone. The key technical contributions of our work are (i) an improved algorithm for estimating the gravity component of the accelerometer measurements, (ii) a novel class of features, extracted from the horizontal accelerometer representation, that are capable of capturing characteristics of acceleration and breaking patterns for different motorised transportation modalities; and (iii) a hierarchical decomposition of the overall detection task. As our experiments demonstrate, the combination of these contributions provides significant improvements in the accuracy of transportation mode detection, in particular for motorised transportation modalities. We have implemented our approach on Android smartphones and integrated it as part of a mobile application that aims at motivating people to reduce their CO2 consumption [13]. In the remainder of this section we describe the different components of our approach in detail.

3.1 Overview

Our approach decomposes transportation mode detection hierarchically into subtasks, proceeding from a coarse-grained classification towards a fine-grained distinction of transportation modality. At the core of our system are three classifiers, which are organized into a hierarchy; see Fig. 1. At the root of the hierarchy is a *kinematic motion classifier* which performs a coarse-grained distinction between pedestrian and other modalities. When the kinematic motion classifier fails to detect substantial physical movement, the process progresses to a *stationary* classifier, which determines whether the user is stationary or in a motorised transport. When motorised transportation is detected, the classification proceeds to a *motorised* classifier which is responsible for classifying the current transportation activity into one of five modalities: bus, train, metro, tram or car.

Changes in transportation behavior typically occur infrequently and each activity has duration of several minutes. Furthermore, changes from one motorised transportation modality to another are typically separated by easily detectable walking segments [36]. Our approach treats the different transportation activities as segments instead of
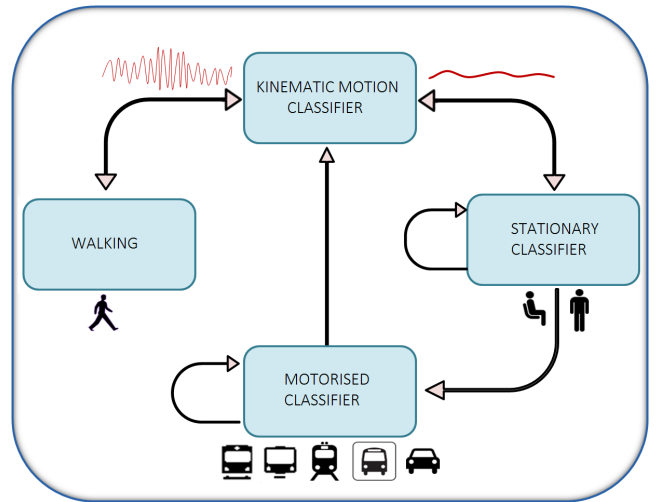

Figure 1: Overview of the classifiers used in our system and their dependencies.

performing solely frame-by-frame classification. Segments contain more information than the frames can express individually, resulting in improved classification performance. Specifically, as more evidence to support one of the modalities accumulates during a segment, the prediction becomes increasingly accurate. The segment-wise classification is continued until change in transportation mode is detected. Stationary periods within motorised modality are interpreted as being in a stopped vehicle, e.g., due to traffic lights or stopping at a station. A detailed description of the frame and segment-based classifiers are is given in Sec. 3.4.

3.2 Preprocessing and Gravity Estimation

We consider three dimensional acceleration measurements obtained from contemporary smartphones. We preprocess the raw measurements by applying a low-pass filter that retains 90% of energy. This is performed to remove jitter from the measurements and is in line with current best practices. Next, we aggregate the measurements using a sliding window with 50% overlap and a duration of 1.2 seconds. The length of the window was selected to ensure the monitoring can rapidly react to changes in the transportation behavior of the user. Once the measurements have been filtered, we project the sensor measurements to a global reference frame by estimating the gravity component along each axis and calculating gravity eliminated projections of vertical and horizontal acceleration. We consider a novel method for estimating the gravity component from accelerometer measurements that improves the robustness of gravity estimation, particularly in the presence of sustained acceleration.

Currently the dominant approach for estimating the gravity component from accelerometer measurements is to use the mean over a window of fixed duration [19, 22]. While elegant and simple, this approach, first proposed by Mizell [22], suffers from two fundamental limitations. First, this approach is inherently based on the assumption that, given a sufficiently long window of measurements, noise and observed accelerometer patterns are uncorrelated over time. This assumption does not hold during sustained accelera-

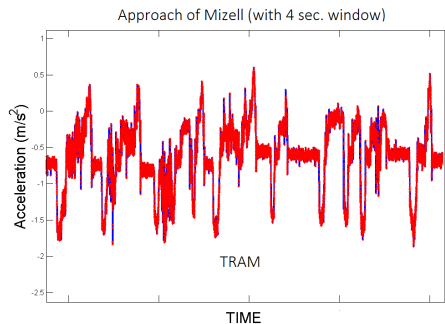**Algorithm 1** Gravity $(Accelerometer_{window}, TH_{var})$

1: $W_{mean} = mean(Accelerometer_{window})$
2: $W_{var} = var(Accelerometer_{window})$
3: **if** $||W_{mean} - G_{est}|| \geq 2\text{m/s}^2$ **then**
4:     $TH_{var} = \epsilon$ ▷ Reset variance threshold
5: **end if**
6: **if** $W_{var} < 1.5$ **then**
7:     **if** $W_{var} < TH_{var}$ **then**
8:         $G_{est} = W_{mean}$
9:         $TH_{var} = (W_{var} + TH_{var})/2$
10:         $VarIncrease = TH_{var} * \epsilon_{inc}$
11:     **else**
12:         $TH_{var} = TH_{var} + VarIncrease$
13:     **end if**
14: **else**
15:     $G_{est} = \texttt{MizellEstimate(5s)}$
16: **end if**

(a) Mizell with a four second window.

(b) Our approach.

Figure 2: Comparison of the gravity estimation between the algorithm of Mizell and our approach. The estimated gravity corresponds to the solid red line.

	Our	Miz-1	Miz-10	Miz-30
Bus	0.32	0.09	0.24	0.27
Train	0.54	0.16	0.25	0.42
Metro	0.51	0.18	0.32	0.41
Tram	0.35	0.09	0.31	0.32

Table 1: Correlation coefficient between integral of the horizontal gravity eliminated acceleration and GPS speed using our method versus using Mizell's method.
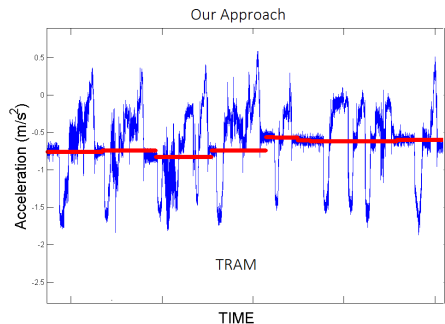
tion, e.g., during any motorised transportation. Second, when the orientation of the sensor suddenly changes, e.g., when the user sits down or stands up, there is a considerable lag before the gravity estimates are accurate again. The lag of the approach can be reduced by shortening the time window over which gravity is estimated, as has been used in several sensing systems [19, 35]. However, this improvement comes with a decrease in the accuracy of the gravity estimates, making it difficult to detect sustained acceleration. To illustrate these limitations, Figure 2(a) shows the gravity estimation produced during a tram ride by the approach used in the Jigsaw system [19], i.e., by using the mean over a four second window. From the figure we can observe that this approach tracks the raw acceleration measurements too closely, removing all information that is relevant for distinguishing between the different transportation modalities.

To overcome the above mentioned limitations, we have developed a novel algorithm for estimating the gravity component of accelerometer measurements. Our approach, summarized in Alg. 1, considers short data windows and estimates the gravity by opportunistically identifying periods where the variation in sensor measurements is sufficiently small, i.e., below a suitable threshold. During these periods, the sensor is approximately stationary, which means that the main force exerting the sensor values is gravity. In many situations, such as walking, bicycling, or traveling with motorised transportation along an uneven road, the measurements contain large variation for a sustained period of time and no opportunities for gravity estimation occur. To estimate gravity during these situations, we dynamically adjust the variance threshold according to the current movement patterns. We allow the variance threshold to increase until a hard upper threshold is reached (currently, we use variance of 1.5), after which the gravity estimates would become overly inaccurate, and utilizing Mizell's technique becomes more suitable.

To reduce the influence of orientation changes on the gravity estimate, we reset the estimate of the gravity component when a large shift in orientation is observed. These shifts are typically caused by extraneous activities, such as user interaction or shifts in orientation due to, e.g., standing up or sitting down. We detect shifts in orientation by comparing the current gravity estimate against the mean of the last measurement window. Whenever these differ by more than

a specific threshold (currently, we use 2m/s) along any of the axes, we re-initialize the gravity estimate for each axis to the mean of current accelerometer window. As illustrated in Fig. 2(b), our gravity estimation is particularly effective while traveling within motorised transportation where periods of low variance typically are interleaved within acceleration and deceleration patterns. In the scenario depicted in the figure, the phone's orientation was relatively stable throughout the tram ride, implying that the estimated gravity component should be approximately constant.

To further demonstrate the benefits of our gravity estimation algorithm, we have conducted a small-scale experiment using a dataset consisting of slightly over 7 hours or data from different motorised transportation modalities. In this experiment, we have compared speed information obtained from GPS with the numeric integration of the gravity eliminated horizontal projection of the acceleration. The numeric integration calculates the area under the gravity eliminated

horizontal acceleration, which can be used to estimate the speed of the user [6]. The results of this experiment, presented in Table 1, demonstrate a strong correlation between the gravity eliminated horizontal projection and the speed information obtained from the GPS. Compared to the approach of Mizell, our algorithm provides better correlation with the speed obtained from GPS for all the evaluated cases, despite using a short time window.

3.3 Feature Extraction

Once the sensor values have been preprocessed and transformed, we construct gravity eliminated horizontal and vertical representations of the accelerometer measurements. We extract features on three different levels of granularity. The three sets of features are referred to as frame-based, peak-based and segment-based features with respect to the feature source. Below we detail each feature set and describe their function in the detection task. For a full list of features; see Table 2.

Frame-based features

The frame-based features considered in our study were chosen based on an analysis of accelerometer features conducted by Figo et al. [6]. From each frame, we extract 27 features from both vertical and horizontal representations, i.e., the total number of features we consider from each frame is 54. The features we extract include statistical features (e.g., mean, variance and kurtosis), time-domain metrics (e.g., double integral, auto-correlation and zero crossings) and frequency-domain metrics (e.g., energy, six first FFT components, entropy and the sum of FFT coefficients). The frame-based features are able to effectively capture characteristics of high-frequency motion caused by, e.g., user's physical movement during pedestrian activity, or during motorised periods, from vehicle's engine and contact between its wheels and surface.

Peak-based features

While the frame-based features can effectively capture information from high-frequency motion, they are unable to capture movement with lower frequencies, such as acceleration and breaking periods of motorised vehicles, which are essential for distinguishing between the different motorised transportation modalities. To capture features from these key periods of vehicular movement, we use the horizontal acceleration projection to extract a set of novel peak-based features that characterize acceleration and deceleration periods. As the kinematic activities are largely characterized by high-frequency motion, we extract the peak-based features only during stationary and motorised periods, i.e., when the kinematic classifier fails to detect substantial, cyclic kinematic movement.

To extract these features, we identify so-called *peak areas* that correspond to acceleration or breaking events; see Figure 3. We identify peak areas by first applying a stream-based event detection algorithm to identify significant changes in the gravity eliminated horizontal acceleration. 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 gravity eliminated horizontal acceleration is observed, i.e., until the horizontal acceleration levels out. Currently we use a pre-defined threshold of

Domain	Features
Statistical	Mean, STD, Variance, Median, Min, Max, Range, Interquartile range Kurtosis, Skewness, RMS
Time	Integral, Double integral, Auto-Correlation, Mean-Crossing Rate,
Frequency	FFT DC,1,2,3,4,5,6 Hz, Spectral Energy, Spectral Entropy, Spectrum peak position, Wavelet Entropy, Wavelet Magnitude
Peak	Volume (AuC), Intensity, Length, Kurtosis, Skewness
Segment	Variance of peak features (10 features), Peak frequency (2 features), Stationary duration, Stationary frequency

Table 2: Full list of the features considered for our classifiers.
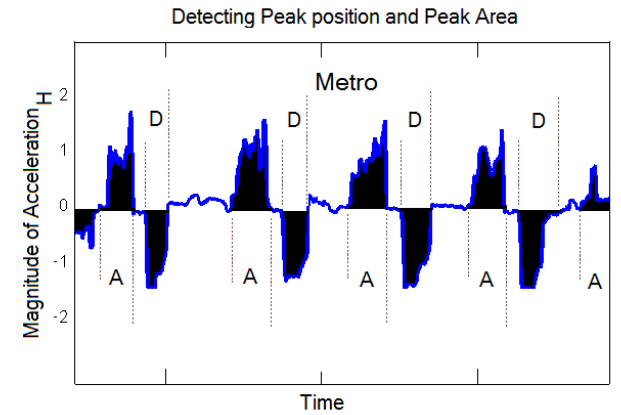

Figure 3: Peak areas detected from gravity eliminated horizontal acceleration during a metro ride.

$0.2m/s^2$ as the threshold for identifying the end boundary of the peak area. Once the starting and ending boundaries have been identified, we extract a set of statistical features that characterize the peak area; see Table 2 for the features that are considered. We calculate these features separately for peaks corresponding to acceleration and to breaking periods, resulting in 10 peak features.

Segment-based features

In addition to the frame and peak-based features, we extract segment-based features that characterize patterns of acceleration and deceleration periods over the observed segment, i.e., during a period of stationary or motorised movement. The segment-based features we consider are the frequency of acceleration and breaking periods, the frequency and duration of the intermittent stationary periods, and the variance of individual peak-based features. The former two of these are analogous to the velocity change rate and stopping rate features that Zheng et al. [36, 37] use as part of their GPS-based transportation mode detection approach. In total we consider 14 segment-based features.

To illustrate the potential of the peak and segment-based features to provide fine-grained detection of motorised trans-
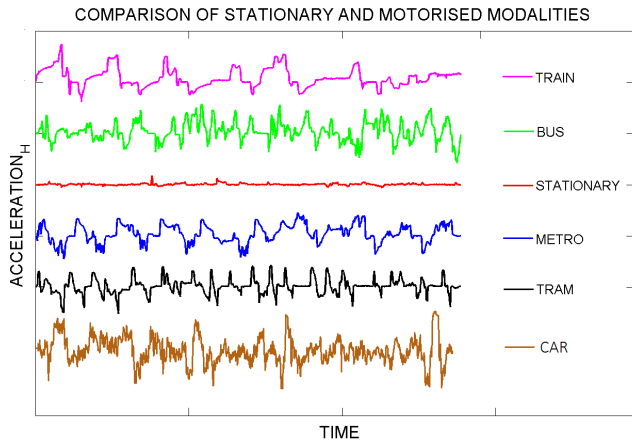
Figure 4: Comparison of the horizontal acceleration profiles for different motorised transportation modalities.
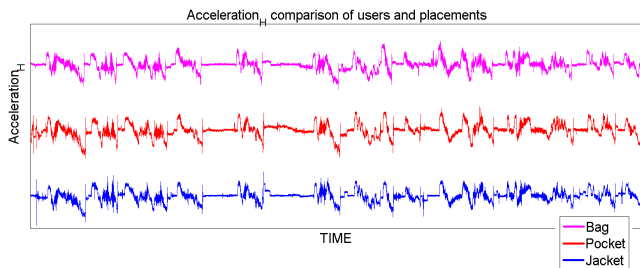

Figure 5: Unfiltered horizontal acceleration during a tram ride for three different users, with phones placed in three different placements.

portation modalities, Fig 4 illustrates the gravity eliminated horizontal acceleration profiles for different motorised transportation modalities. From the figure we can observe that the shape of the peak areas is similar within the same modality, but also distinctive across the different modalities.

The performance of smartphone-based activity recognition approaches is typically sensitive to the placement of the device. Our peak and segment-based features describe the movement patterns of vehicles, instead of those of the user, thus making these features robust over different device placements. To demonstrate this aspect, consider Figure 5, which presents horizontal accelerometer measurements during a single tram ride from three different users and three different phone placements: bag (top, magenta), trouser pocket (middle, red) and jacket (bottom, blue). From the figure, we can observe that the horizontal acceleration profiles are nearly identical, demonstrating that the peak and segment-based features are not susceptible to device placement; see also the generalization capability evaluation in Section 4.2.

In summary, we extract 54 frame features, 10 peak features and 14 segment features, resulting in an overall feature space consisting of 78 features.

3.4 Classification

As discussed in Sec. 3.1, we rely on a three-stage hierarchical classification framework for transportation mode detection. For the kinematic motion classifier, we use a well-established technique [26] of combining an instance-based classifier with a generative classifier, i.e., a discrete Hidden Markov Model (HMM). The stationary and the motorised classifiers, on the other hand, perform segment-based classification where a simple voting scheme is used to aggregate frame-based classifications over the observed segment. The frame-based classifications are obtained using an instance-based classifier. Each of the three classifiers considers a variant of AdaBoost as the instance-based classifier. During our development process we have also evaluated the performance of other instance-based classifiers, including decision trees and support vector machines. Since these techniques were less accurate than AdaBoost, further discussion about these techniques is omitted. In the following we first briefly introduce AdaBoost, after which we detail the segment-based classification and the classifiers that are part of our system.

3.4.1 Adaptive Boosting

Boosting is a general technique for improving the accuracy of any learning algorithm [7, 8]. The basic idea in boosting is to iteratively learn weak classifiers that focus on different subsets of the training data and to combine these classifiers into one strong classifier. Adaptive boosting (AdaBoost), introduced by Freund and Schapire in [7], extends the idea of boosting by tuning to the problematic samples misclassified by previous classifiers. Specifically, AdaBoost operates by assigning each sample in the training data a weight that determines the importance of the sample. Over a series of rounds $t = 1, \ldots, T$, classifiers that minimize classification error on the weighted data are learned. After each round, the weights of the samples are re-evaluated to assign higher priority to samples that are misclassified. Note that the features that are selected by the weak learners provide us with an automatic way to identify the most relevant features for the final classifier design. In our case we use decision trees with depth of one or two as the weak classifiers. The number of boosting rounds T was determined using the scree-criterion, i.e., we plot the classification error for varying values of T and select a suitable value of T, which balances between the classifier accuracy and classifier complexity. In order to retain classifier simplicity, we opted for the minimal T value, after which further increasing the value of T resulted only in marginal gain in accuracy.

3.4.2 Segment-Based Classification

Within each non-pedestrian segment, we assume that the transportation modality remains unchanged. This follows a logical assumption that transition between consecutive stationary or motorised transportation modes require an intermittent period of pedestrian activity [36]. The transportation modality of the segment is predicted based on the entire history of the observed segment using two sources of information:

1. We aggregate classification results of frame,- and peak-based features over the observed segment. As AdaBoost performs soft classification, i.e., it outputs a numeric value indicating the voting result of the weak classifiers, we use a mean value of the voting results over the duration of the segment.

Representation	Features
Horizontal	Variance, Range,
Vertical	Variance, Interquartile Range
	FFT 2Hz

Table 3: Features used in the kinematic motion classifier.

Representation	Features
Horizontal	Range, Root Mean Square,
	Entropy, Max Coeff, Variance
Vertical	Variance, Max Coeff, Range,
	Root Mean Square, FFT 2,3,5Hz
Peaks	Volume, Frequency, Intensity, Length
Segment	Variation of Peak Lengths

Table 5: Features used in the motorised classifier.

2. We compute the classification result of the segment-based features over the observed segment.

After acquiring classification results from the two information sources, we obtain the final classification by combining the results of the two classifier outputs. To retain simplicity of the system, we currently simply use the average of the two sources.

3.4.3 Kinematic Motion Classifier

The kinematic motion classifier utilizes the frame-based accelerometer features extracted from each window to distinguish between pedestrian and other modalities. For the classification, we use decision trees with depth one (i.e., decision stumps) and a combination of 10 weak classifiers, i.e., $T = 10$. The features that were selected for the final classifier design are listed in Table 3. Clearly the most effective features for this task are the variation for both horizontal and vertical representations. Analogous to the variance, also range and interquartile range were selected. This is due to high variation of pedestrian modes compared to stationary or motorised transportation modes. Another effective feature selected by the AdaBoost algorithm is the FFT 2Hz component, which captures the repetitive nature of walking, typically with $1 - 3s$ interval. The accuracy of the kinematic motion classifier is over 99% which means that it can robustly determine the subsequent classifier to apply.

3.4.4 Stationary Classifier

Representation	Features
Horizontal	Variance, Root Mean Square,
	Max Coeff, Wavelet Entropy
Vertical	Interquartile Range, Variance,
	Root Mean Square,
	FFT 2Hz
Peaks	Volume, Length, Frequency
Segment	Variation of Peak Lengths

Table 4: Features used in the stationary classifier.

The stationary classifier uses both the peak features and the frame-based features for distinguishing between stationary and motorised periods. The best performance was obtained using a combination of $T = 15$ weak learners, each comprising a decision trees of depth two. The list of selected features is shown in Table 4. The key features for this task are the volume, length and frequency of acceleration and breaking patterns. The frame-based features that are selected primarily characterize the intensity of movement along the horizontal and vertical axes. These features include the root mean square, variance and range. In addition, features characterizing the pattern of movements, i.e.,

FFT 2Hz, Maximum coefficient and Wavelet Entropy, help to distinguish between the more repetitive movement pattern of motorised transportation from sporadic movement within stationary periods. The complexity of this classifier indicates the difficulty of achieving this task using only the kinematic information. Nevertheless, the precision of the final classifier is over 95%, which enables our approach to provide robust continuous detection of transportation modality even in the presence of periods where stationarity is interleaved with motorised transportation.

3.4.5 Motorised Classifier

The motorised classifier is responsible for distinguishing between different motorised transportation modalities. Currently the classifier supports distinguishing between car, bus, train, tram and metro. As movement patterns between these different transportation modalities are very similar, this is the most challenging task in our design. For the AdaBoost classifier we use $T = 20$, and decision trees of depth two as the weak learners. The final set of selected features is shown in Table 5. The frequency of the acceleration and breaking peaks are the most important features as they enable distinguishing between vehicles that move alongside other traffic (i.e., car, bus and tram) and vehicles moving independently of other traffic (i.e., train and metro). The intensity and length of the acceleration and breaking periods are effective in separating car from other motorised modalities, as the quicker driving maneuvers make car clearly distinguishable from larger vehicles. The intensity and volume of acceleration and breaking periods, on the other hand, are effective in separating the slow moving tram from other motorised modalities.

The motorised classifier uses frame-based features primarily to distinguish vehicles that operate on roads from vehicles operating on rails. The selected frame-based features effectively capture characteristics of vertical movement as well as overall noisiness of the measurements. However, compared to the other two classifiers, the frame-based features are assigned significantly smaller weight, indicating a low relevance to classification. As our experiments demonstrate, our design can effectively detect the correct modality with approximately 80% precision. Moreover, most of the classification errors occur during the first minutes of each motorised segment before sufficient information about acceleration and breaking patterns has accumulated. With our current design, the most problematic task is to distinguish between metro and commuter train as both have very similar frame-based and peak features. These errors could be alleviated, e.g., by fusing in information about transportation routes, or by using features from other sensors; see Sec. 5.

4. EVALUATION

4.1 Datasets

To develop and evaluate our approach, we have collected over 150 hours of transportation data from 16 individuals and 4 different countries. In the data collection we have considered three smartphone models: Samsung Nexus S, Samsung Galaxy S II and Samsung Galaxy S III. On the Nexus S, the accelerometer is sampled at 60Hz frequency, whereas on the Galaxy S II and Galaxy S III smartphones the accelerometer is sampled at 100Hz frequency. The measurements that we have collected consist of four different datasets:

- Data collected from a predefined Scenario A during winter of $2011-2012$. Total of approximately 44 hours of data was collected by nine individuals.

- Data collected from a predefined Scenario B during autumn of 2012. Total of approximately 23 hours of data was collected by seven individuals.

- Data collected from everyday transportation in Finland between late 2011 and late 2012. Total of 65 hours of data was collected by seven individuals.

- Transportation data collected from abroad between 2011 and 2012. Total of over 20 hours of transportation data was collected by three individuals.

The two scenarios used in the data collection took place within the downtown area of Helsinki, Finland. The scenarios were designed to contain a large variety of everyday transportation modes and to last between 90 and 120 minutes per participant. The first scenario was carried out during the winter of $2011-2012$ and included nine participants. The second scenario was carried out during autumn of 2012 and included seven participants. The scenarios were designed to contain complementary sets of participants, public transportation lines and traffic conditions; see Fig 6(a) and Fig 6(b) for an illustration of the data collection scenarios. To ensure that the results of the experiments are not sensitive to the placement of the sensor, the scenario data was collected from different sensor placements. We considered the three most common placements for a mobile phone in an urban space [11]: trouser pockets, bag and jacket pockets. To avoid disturbing the sensing units that were used to collect sensor values, the ground truth annotations were made on an additional mobile phone that was not part of any of the sensing units collecting sensor values.

In addition to the data collection scenarios, we have collected everyday transportation data from seven individuals. The participants were asked to collect sensor data during their everyday transportation and to record ground truth labels using our data collection application. No constraints on the phone placement or usage were placed on the user. The data was collected with one of the three phones models at our disposal, Nexus S, Galaxy S II or Galaxy S III. The everyday data covers a wide range of transportation behavior within our target city; a total of over 30 different routes were traveled during various times and traffic conditions.

To assess the generalization capability of our approach, we have also collected data from various abroad destinations: Japan (Tokyo, Amnon, Shirakami, Hiroshaki), Germany (Frankfurt, Saarbrüken), and Luxembourg. This data was collected by three individuals, following the same approach as with the everyday data collection. The measurements were collected using Galaxy S II and Galaxy S III phones.

4.2 Performance Evaluation

We consider the following aspects in our evaluation: accuracy of transportation mode detection, power consumption, generalization performance of the classifiers, and the robustness and latency of the detection. The accuracy of the classifiers was measured using standard classification metrics, whereas robustness and latency were measured using the event and frame-based metrics proposed by Ward et al. [32, 33]. As part of the evaluation, we compare our system against the accelerometer-based approach of Wang et al. [31] and the approach of Reddy et al. [26], which uses both GPS and accelerometer measurements.

The approach of Wang et al. represents a typical accelerometer based transportation mode detection system, whereas the approach of Reddy et al. represents current state-of-the-art in smartphone-based transportation mode detection. We consider the following seven classes in our evaluation: Stationary, Walk, Bus, Train, Metro, Tram and Car. We omit fine-grained classification of pedestrian modalities, focusing instead on fine-grained detection of motorised transportation modalities as previous work has already demonstrated high, over 90% accuracies [26, 1] for the different pedestrian modalities.

Detection Accuracy

We first consider the capability of our approach to detect different transportation modalities. To carry out this evaluation, we consider all of the data collected from the two scenarios and evaluate detection accuracy using leave-one-user-out cross-validation. The evaluation of car modality is postponed until the evaluation of generalization performance as it was not practically feasible for us to collect driving measurements from the two scenarios.

The results of this evaluation together with a comparison against the approaches of Wang et al. and Reddy et al. are shown in Table 6. The mean precision and recall of our approach is over 80%, demonstrating that it can accurately distinguish between different transportation modalities despite relying solely on accelerometer measurements. Moreover, the variance of the results is relatively small, demonstrating robustness across users. Compared to the the two baseline systems, our approach provides over 20% higher precision and recall than the approach of Wang et al. and over 10% higher precision and recall than the approach of Reddy et al. The most notable performance differences can be observed for motorised and stationary periods. We assessed the statistical significance of the performance differences using one-sided t-tests. The results of the significance tests indicate that our approach is significantly better than the approach of Wang et al. for all modalities except bus. Compared to the approach of Reddy et al., our approach provides significantly better precision for stationary, bus and train, and significantly better recall for train, metro and tram.

With respect to the baselines, the results for the approach of Wang et al. are analogous to the those reported by the authors. However, the mean precision and recall of the approach of Reddy et al. appear significantly lower than reported by the authors. There are two reasons for this differ-
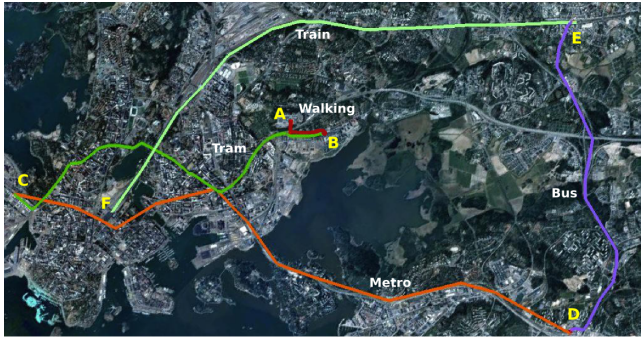
(a) Scenario A
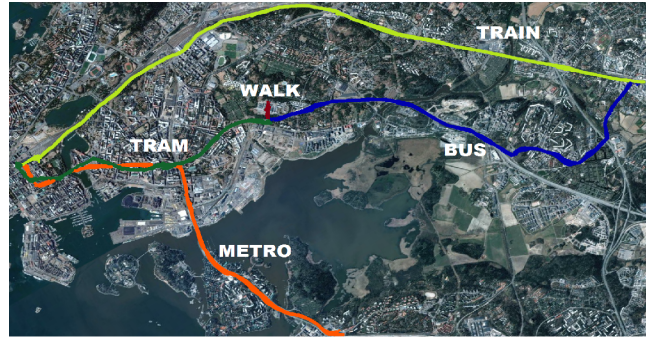
(b) Scenario B

Figure 6: Overview of the scenarios which were used for data collection.

	Precision			Recall		
	Peaks	Wang	Reddy	Peaks	Wang	Reddy
Stationary	96.1 (0.5)	57.3 (4.5)	81.6 (1.0)	70.0 (2.1)	59.5 (2.3)	70.6 (2.9)
Walk	93.1 (0.1)	87.2 (0.2)	97.7 (0.1)	95.9 (0.1)	89.1 (0.2)	95.9 (0.1)
Bus	78.2 (4.2)	71.1 (1.4)	67.3 (1.6)	78.0 (3.3)	70.4 (1.4)	86.2 (6.4)
Train	68.2 (5.0)	32.1 (0.8)	7.7 (4.4)	80.1 (4.0)	31.6 (0.7)	55.4 (11.9)
Metro	64.5 (5.9)	54.4 (0.6)	70.1 (8.8)	82.0 (2.6)	51.4 (0.9)	56.6 (3.5)
Tram	84.0 (2.1)	58.1 (0.8)	82.8 (7.5)	86.1 (2.1)	58.2 (0.8)	64.5 (7.0)
Mean	80.1 (2.9)	60.0 (1.4)	68.0 (3.9)	82.1 (2.4)	60.2 (1.1)	71.6 (5.3)

Table 6: Detection accuracy and variance for our system, and a comparison against the baseline systems.

ence. First, Reddy et al. merged all motorised transportation modalities into the same class, whereas we consider a fine-grained distinction of different modalities. In case we aggregate all motorised modalities into the same class, the approach of Reddy et al. achieves a precision of 92.2% and a recall of 96.1% which is analogous to the results reported by the authors. The second reason for the performance difference is that the stationary class essentially corresponds to a null-class, which suffers from an ambiguous definition. Examples of situations that would fall within the stationary class include (i) leaving the phone idle, e.g., on table, (ii) having a dinner, working or otherwise sitting for prolonged period of time, and, (iii) short stationary periods within motorised transportation activity. The amount of kinematic movement contained in each of these cases differs significantly, resulting in varying results. In our case, the stationary class contains data also from periods that are interleaved within active transportation, i.e., while waiting for the next transportation vehicle or intermittent stops within pedestrian activity due to, e.g., traffic lights or queuing to enter a vehicle. As these periods naturally contain some movement, this creates stationary data which closely resembles the jitters and vibrations of motorised transportation, rendering these modalities hard to separate with solely frame-based features. Moreover, our dataset contains several stationary periods, where reliable speed information from the GPS cannot be polled, e.g., while waiting for metro underground or while waiting for train or bus within a station.

To better characterize the performance of our approach, Table 7 presents the confusion matrix of the evaluation for our approach. From the confusion matrix we can observe that the most challenging individual subtask is to distinguish between train and metro. The difficulty of this task

Stationary	Walk	Bus	Train	Metro	Tram
62733	755	135	58	302	361
1549	63664	456	650	723	805
3976	647	32400	18	104	8118
6730	330	874	31921	5907	1894
5057	711	2961	10879	41203	2682
4341	318	10123	87	2067	77715

Table 7: Confusion matrix for our approach.

TMode	Precision	Recall
Stationary	61.9 (-34.2)	64.0 (-6.0)
Walk	93.0 (-0.1)	93.0 (-2.9)
Bus	71.6 (-6.6)	71.5 (-6.5)
Train	25.1 (-43.1)	54.9 (-25.2)
Metro	60.1 (-4.4)	56.0 (-26.0)
Tram	69.6 (-14.4)	66.7 (-19.4)
Mean	63.6 (-16.5)	67.7 (-14.3)

Table 8: Detection accuracy for cross-user evaluation without the peak features.

is also evident from the results of the baseline approaches. The difficulty is caused by the close similarity of the peak, segment and frame features for these modalities, as also discussed in Sec. 3.4.5. In case these classes are combined into the same class, the classification accuracy improves significantly, reaching almost 90% precision.

To evaluate the effectiveness of the peak and segment-based features, we have repeated the experiment using our

TMode	Precision	Recall
Stationary	95.1 (2.0)	72.4 (2.1)
Walk	92.7 (2.4)	92.4 (2.0)
Bus	84.8 (3.6)	79.6 (0.5)
Train	74.7 (4.0)	80.8 (1.9)
Metro	69.6 (3.8)	81.3 (1.7)
Tram	88.4 (2.5)	87.9 (1.6)
Mean	84.2 (3.0)	82.4 (1.6)

Table 9: Results of cross-placement evaluation.

approach with only the frame-based feature. The results of this evaluation are shown in Table 8 and show a significant decrease in detection accuracies, in particular for the motorised modalities and stationary periods. Consequently, the novel peak and segment-based features considered in our work are effective at improving detection performance. The accuracies resulting from the use of frame-based features only, however, still exceed those of Wang et al., suggesting that our gravity-estimation algorithm is able to provide more expressive features than the L_2 norm used in the approach of Wang et al.

Influence of Device Placement

A key challenge in mobile activity recognition research is how to deal with variations in sensor placement effectively. To demonstrate that our approach is robust against device placement, we carried out the scenario evaluation described previously also using leave-one-placement-out cross-validation. The results of this evaluation are shown in Table 9. The detection performance for most activities is higher than in the cross-user case, indicating that our approach is robust against variations in device placement. Accordingly, variations in user characteristics have a more significant influence on the results than variations in placement, however, generally our approach is robust against both sources.

Power Consumption

Any real-world deployment of continuous transportation mode detection on smartphones requires that the detection has minimal impact on the operational time of the smartphone, i.e., that the detection's power consumption is minimal. We have performed a coarse-grained evaluation of the power consumption using empirical power models that were constructed from measurements obtained using the Monsoon Power Monitor[1]. To construct the power models, we first constructed a model of the background power consumption of the device. The power consumption values of the different approaches were then obtained by measuring overall power consumption and subtracting the background consumption.

The average power consumption of our proposed system is 85 mW. From the total power consumption 21 mW results from accelerometer sampling, 24 mW from constructing the accelerometer representations, and 35 mW from calculating the frame-based features. The remaining 5 mW results from peak and segment feature calculations. Compared to the power consumption of the system presented by Reddy et al., our approach provides a 155 mW reduction in energy

[1]http://www.msoon.com/LabEquipment/PowerMonitor/ [Retrieved: 2012-12-09]

Application	Energy
Peaks TMD	**85 mW**
Wang TMD	50 mW
Reddy TMD	240 mW
Active call	680 mW
Music player	50 mW
Video recording	930 mW
Video playing	660 mW
Accelerometer	21 mW
Magnetometer	48 mW
Gyroscope	130 mW
Microphone	105 mW
GPS sampling	176 mW
Background	140 mW
Phone screen	470 mW

Table 10: Energy consumption of our system, along with a comparison with the baseline systems and a selection of popular applications.

consumption. Compared to the system presented by Wang et at., our approach consumes 35 mW more energy due to more complex accelerometer processing. For a comparison between baselines systems and other popular applications see Table 10. The numbers presented above do not take into account duty-cycling or other methods for reducing sensor or CPU energy consumption. The inertial sensors, i.e., accelerometer, magnetometer and gyroscope were sampled at maximal sampling rate. The GPS sensor was sampled similar to the baseline system at 1 Hz frequency.

The real-world implication of the presented energy comparison is that systems based on heavy use of GPS or other power hungry sensors can deplete the battery of a smartphone within hours, while systems based on accelerometer, such as ours, can potentially run in the background of a smartphone throughout the entire day.

Generalization Performance

We next demonstrate that our proposed approach generalize over various routes, vehicle types, traffic conditions and geographic locations. To demonstrate this aspect, we consider an experiment where all classifiers are trained using the data from Scenario A and B and evaluated using Everyday data and data collected from abroad. To evaluate the detection accuracy of driving a car, we augmented the training data with driving measurements from the everyday data. Specifically, we included 40% of all driving measurements in the training data. As the test data we consider all remaining data collected from the everyday transportation activities as well as the data collected from abroad.

The results of this evaluation are shown in Table 11. From the results, we can observe that the precision and recall of our approach are similar or even better than those achieved in the controlled experiment. The detection accuracy of car, which was not considered in the scenario-based evaluation, is over 90%, indicating that driving a car is the easiest motorized modality to recognize.

The detection accuracy for stationary, walking and tram are similar or marginally better than those achieved in the controlled experiments. The largest increase is evident for bus, which has increased some 8% for both precision and re-

TMode	Precision			Recall		
	Peaks	Wang	Reddy	Peaks	Wang	Reddy
Stationary	96.0	51.5	80.9	72.9	52.6	78.0
Walk	92.4	84.1	97.7	97.3	85.4	91.1
Bus	85.2	59.1	63.1	86.7	77.3	78.7
Train	75.9	24.8	4.4	80.7	49.3	43.6
Metro	67.1	50.4	58.1	72.7	37.9	35.3
Tram	87.7	70.9	72.4	90.0	42.1	40.1
Car	90.1	79.3	89.9	96.7	80.1	95.4
Mean	84.9	60.0	66.7	85.3	60.7	66.0

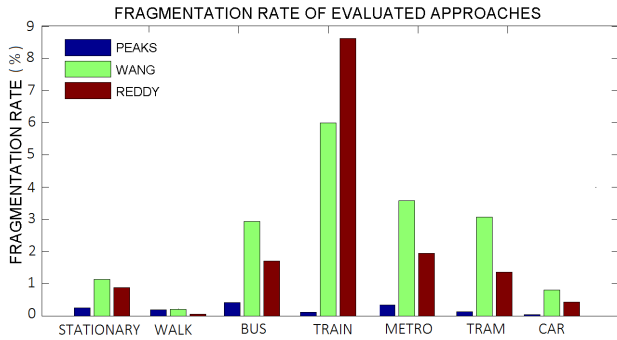Table 11: **Generalization experiment of our detection system.**

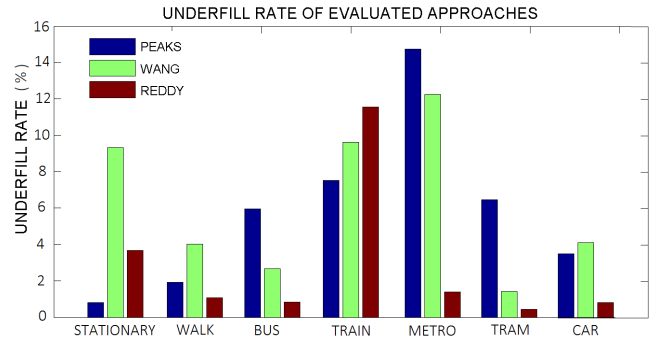Figure 7: **Fragmentation rate of the detection performance.**

Figure 8: **Underfill Rate of the detection performance.**

call. This is due to the long-distance bus trips present in the everyday dataset, which suit our segment classification particularly well. For similar reasons, the precision of train has increased some 7% from the controlled experiments. The accuracy increase is moderated by the different characteristics of the vehicles used abroad, in particular the bullet train, which was particularly challenging for our classifiers which were trained with data from commuter trains in Helsinki. Metro, while increased 3% for precision, has dropped some 10% for recall. This is mainly caused by mixed classification between train and metro for commuter trains in Tokyo, as these represent qualities from both, train and metro, and are hard to label accurately. For tram, both precision and recall have increased by 3%.

In comparison to our approach, the performance of both baselines decrease from the first evaluation, demonstrating that our system generalizes well and is better at capturing generic characteristics of different transportation modalities. As the test data contains measurements collected in countries with significantly different transportation infrastructures and spatial topologies, the results also effectively demonstrate that our approach generalizes well across different geographic locations.

Detection Robustness

The evaluation of the transportation modality detection gives an incomplete view of the suitability of our approach for applications that continually monitor the transportation modality of the user. In particular, the classification metrics considered in the previous section ignore the continuous nature of the transportation activities. The metrics give no indica-

tion about the robustness of the detection and the latency between detecting changes in transportation mode.

To evaluate detection robustness, we consider the fragmentation rate of the different activities, i.e., the fraction of all ground truth events that the classification recognizes as multiple events. The results of this evaluation and a comparison against the baseline approaches are shown in Fig. 7. With the exception of walking modality, for which the fragmentation rate is under 2.0% for all evaluated systems, our approach provides significantly more robust transportation modality detection. On average, the fragmentation rate of our approach is less than one-tenth compared to the baselines. The increased robustness is mainly due to the segment classification used in our approach.

As part of the robustness analysis, we have also examined the latency of detecting changes in transportation mode. To illustrate this aspect, Fig. 8 shows the underfill rate of the detection at the start of a new modality. The underfill rate effectively measures the rate at which there is a delay in detecting the correct modality when the modality has changed. From the results we can observe that both our approach and the approach of Reddy et al. are effective at detecting changes from or to pedestrian transportation behavior. The approach of Wang et al. uses a longer window (i.e., 8 seconds), which increases the latency in detecting changes between transportation modes.

For motorised modalities, our system has higher latency than the system of Reddy et al. The difference is a result from the use of peak and segment-based features, which require observing few peaks before they can accurately detect the transition from stationary to motorized transportation,
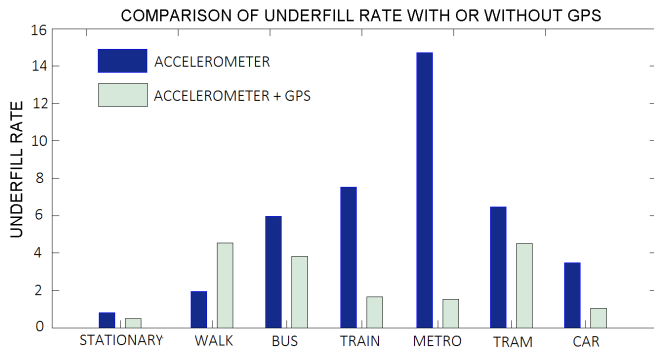
Figure 9: The effect on underfill rate from using GPS sensor for stationary detection.

and to identify the correct motorized mode. The detection latency can be easily reduced by infusing any sensor capable of stationary detection (i.e., GPS, GSM or WiFI) to our stationary and motorised classifiers when the importance of low latency outweighs the associated energy cost. This aspect is illustrated in Fig. 9, which shows how the inclusion of GPS speed significantly reduces the underfill rate of our approach for all of the motorised transportation modalities. A possible extension of our approach that could reduce detection latency, as well as reduce misclassifications, thus is to sample the GPS whenever a transition from kinematic/pedestrian modality to motorized transportation is observed.

5. DISCUSSION AND SUMMARY

The present paper has introduced novel accelerometer techniques for transportation mode detection on smartphones. Evaluations conducted using over 150 hours of transportation data demonstrate that, compared to existing accelerometer based solutions, our work provides over 20% improvements in the accuracy of transportation mode detection, while at the same time improving robustness of the detection. The results also demonstrate that our approach generalizes well across users and geographic locations. The primary technical contributions of our work are an improved algorithm for estimating the gravity component of accelerometer measurements, a novel set of accelerometer features which characterize acceleration and breaking patterns during motorised transportation, and a hierarchical decomposition of the overall recognition task. We have implemented our system on Android smartphone and are currently using it as part of a mobile application that supports persuasive mobility.

While the present work has focused exclusively on transportation mode detection, our technical contributions have significance to numerous other domains. The possibility to construct accurate reconstructions of the gravity eliminated horizontal and vertical accerelations for vehicular movement open up possibilities, e.g., for recognizing potholes, estimating the skill-level of drivers, recognizing traffic accidents, and so forth. Aggressiveness of braking and acceleration patterns, which can be detected using our peak-based features, are also key factors affecting fuel consumption of drivers, opening up possibilities to build accelerometer-based solutions for assessing the sustainability of driving styles.

Our current system relies solely on the phone's accelerometer, which means that the impact of our system on the phone's battery lifetime is reasonable even when the sensor is polled with maximum sampling frequency. The power consumption of our approach could be easily reduced further. For example, we could use the hierarchical classification architecture to compute only the relevant features for each classification task. The kinematic classifier which relies on computationally light features has the lowest energy consumption, while the motorised classifiers with a wide range of features requires the most energy. As our experiments considered phones with two different sampling rates, the results suggest that our approach is not overly sensitive to the sampling rate of the accelerometer, providing further possibilities to reduce power consumption by duty cycling the sensor, particular during periods of prolonged stationary behavior [34]. During periods of extended stationary behavior, the application could also be switched off until notable kinematic movement is detected from the accelerometer [14] or the primary GSM cell changes [26]. As humans tend to spend most of their time within a limited set of locations, with only occasional transitions between these places [10], significant reductions in power consumption could be achieved by minimizing the power consumption of the stationary classifier.

While our system provides improved robustness compared to other purely accelerometer-based methods, it is still susceptible to interference from extraneous kinematic events, such as user interaction or changes in the orientation of the phone. Consequently, methods for handling these situations, such as those used as part of the Jigsaw system [19], can be used to further improve the robustness of our approach. The current implementation of our approach already integrates some of these techniques, using light and proximity sensors together with information about the phone's state (e.g., ongoing call, screen of, buttons pressed) to identify and discard periods of user interaction.

Currently the main limitation of our approach is the latency of detecting the correct modality while switching to a motorised transportation modality. In applications which require quick detection of correct transportation modality, the latency can be significantly reduced by fusing in measurements from additional sensors, e.g., changes in GSM or Wi-Fi signal environment, GPS speed or changes in magnetic field.

6. REFERENCES

[1] L. Bao and S. S. Intille. Activity recognition from user-annotated acceleration data. In *Proceedings of the 2nd International Conference on Pervasive Computing (PERVASIVE)*, volume 3001 of *Lecture Notes in Computer Science*, pages 1–17. Springer-Verlag, 2004.

[2] G. Bieber, J. Voskamp, and B. Urban. Activity recognition for everyday life on mobile phones. In *Proceedings of the 5th International Conference on Universal Access in Human-Computer Interaction (UAHCI)*, pages 289–296, 2009.

[3] A. Bolbol, T. Cheng, I. Tsapakis, and J. Haworth. Inferring hybrid transportation modes from sparse gps data using a moving window svm classification. *Computers, Environment and Urban Systems*, 31;6:526–537, 2012.

[4] T. Brezmes, J.-L. Gorricho, and J. Cotrina. Activity recognition from accelerometer data on a mobile phone. In *Workshop Proceedings of the 10th International Work-Conference on Artificial Neural Networks (IWANN)*, pages 796–799, 2009.

[5] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, I. Smith, and J. A. Landay. Activity sensing in the wild: a field trial of ubifit garden. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1797–1806, New York, NY, USA, 2008. ACM.

[6] D. Figo, P. Diniz, D. Ferreira, and J. Cardoso. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*, 14-7:645–662, 2010.

[7] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *Proceedings of the Second European Conference on Computational Learning Theory*, 1995.

[8] Y. Freund and R. E. Schapire. A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, 14(5):771–780, 1999.

[9] J. Froehlich, T. Dillahunt, P. Klasnja, J. Mankoff, S. Consolvo, B. Harrison, and J. A. Landay. Ubigreen: investigating a mobile tool for tracking and supporting green transportation habits. In *Proceedings of the 27th international conference on Human factors in computing systems (CHI)*, pages 1043–1052. ACM, 2009.

[10] M. C. González, C. A. Hidalgo, and A.-L. Barabási. Understanding individual human mobility patterns. *Nature*, 453:779–782, 2008.

[11] F. Ichikawa, J. Chipchase, and R. Grignani. Where's the phone? a study of mobile phone location in public spaces. In *Proceedings of the 2nd International Conference on Mobile Technology, Applications and Systems*, pages 1 – 8. IEEE, 2005.

[12] T. Iso and K. Yamazaki. Gait analyzer based on a cell phone with a single three-axis accelerometer. *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pages 141–144, 2006.

[13] A. Jylhä, P. Nurmi, M. Siren, S. Hemminki, and G. Jacucci. Matkahupi: a persuasive mobile application for sustainable mobility. In *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2013.

[14] D. H. Kim, Y. Kim, D. Estrin, and M. B. Srivastava. Sensloc: sensing everyday places and paths using less energy. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, pages 43–56. ACM, 2010.

[15] M. B. Kjærgaard, S. Bhattacharya, H. Blunck, and P. Nurmi. Energy-efficient trajectory tracking for mobile devices. In *Proceedings of the 9th International Conference on Mobile Systems, Applications and Services (MobiSys)*, 2011.

[16] J. Krumm and E. Horvitz. LOCADIO: Inferring motion and location from Wi-Fi signal strengths. In *Proceedings of the 1st International Conference on Mobile and Ubiquitous Systems (Mobiquitous)*, pages 4 – 14. IEEE, 2004.

[17] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell. A survey of mobile phone sensing. *IEEE Communications Magazine*, 48(9):140 –150, 2010.

[18] D. Lazer, A. P. L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. R. 2, and M. V. Alstyne. Computational social science. *Science*, 323(5915):721–723, 2009.

[19] H. Lu, J. Yang, Z. Liu, N. D. Lane, C. T., and C. A. The jigsaw continuous sensing engine for mobile phone applications. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, pages 71–84, 2010.

[20] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell. Sensing meets mobile social networks: the design, implementation and evaluation of the CenceMe application. In *Proceedings of the 6th ACM conference on Embedded network sensor systems (SenSys)*, pages 337–350, New York, NY, USA, 2008. ACM.

[21] T. M. Mitchell. Mining our reality. *Science*, 326(1644):1644–1645, 2009.

[22] D. Mizell. Using gravity to estimate accelerometer orientation. In *Proc. Seventh IEEE International Symposium on Wearable Computers*, pages 252–253, 18–21 Oct. 2005.

[23] M. Mun, D. Estrin, J. Burke, and M. Hansen. Parsimonious mobility classification using gsm and wifi traces. In *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems (SenSys)*, pages 1–5, 2008.

[24] K. Muthukrishnan, M. Lijding, N. Meratnia, and P. Havinga. Sensing Motion Using Spectral and Spatial Analysis of WLAN RSSI. In *Proceedingds of the 2nd European Conference on Smart Sensing and Context (EuroSSC)*, pages 62–76,. Springer, 2007.

[25] P. Nurmi, S. Bhattacharya, and J. Kukkonen. A grid-based algorithm for on-device GSM positioning. In *Proceedings of the 12th International Conference on Ubiquitous Computing (UbiComp)*, pages 227–236, 2010.

[26] S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava. Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks*, 6(2):13:1–13:27, 2010.

[27] T. Sohn, A. Varshavsky, A. LaMarca, M. Y. Chen, T. Choudhury, I. Smith, S. Consolvo, J. Hightower, W. G. Griswold, and E. de Lara. Mobility detection using everyday GSM traces. In *Proceedings of the 8th*

International Conference on Ubiquitous Computing (Ubicomp), pages 212–224, 2006.

[28] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási. Limits of predictability in human mobility. *Science*, 19(5968):1018–1021, 2010.

[29] D. Soper. Is human mobility tracking a good idea? *Communications of the ACM*, 55; 4:35–37, 2012.

[30] L. Stenneth, O. Wolfson, P. S. Yu, and B. Xu. Transportation mode detection using mobile phones and gis information. *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 54–63, 2011.

[31] S. Wang, C. Chen, and J. Ma. Accelerometer based transportation mode recognition on mobile phones. In *Asia-Pacific Conference on Wearable Computing Systems*, pages 44–46, 2010.

[32] J. Ward, P. Lukowicz, and G. Tröster. Evaluating performance in continuous context recognition using event-driven error characterization. In *Proceedings of the 2nd International Workshop on Location- and Context-Awareness (LoCA)*, pages 239–255. Springer, 2006.

[33] J. A. Ward, P. Lukowicz, and H. W. Gellersen. Performance metrics for activity recognition. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(1):6:1–6:23, 2011.

[34] Z. Yan, V. Subbaraju, D. Chakrabarti, A. Misra, and A. K. Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach. *International Symposium on Wearable Computers (ISWC)*, 2012.

[35] J. Yang. Toward physical activity diary: motion recognition using simple acceleration features with mobile phones. In *Proceedings of the 1st International Workshop on Interactive Multimedia for Consumer Electronics (IMCE)*, pages 1–9, 2009.

[36] Y. Zheng, Y. Chen, Q. Li, and W.-Y. Xie, X. Ma. Understanding transportation modes based on gps data for web applications. *ACM Transactions on the Web*, 4,1, 2010.

[37] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma. Understanding mobility based on gps data. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 312–321, 2008.

[38] Y. Zheng, Y. Liu, J. Yuan, and X. Xie. Urban computing with taxicabs. In *Proceedings of the 13th International Conference on Ubiquitous Computing (Ubicomp)*, pages 89–98. ACM, 2011.