

Towards Indoor Transportation Mode Detection using Mobile Sensing

Thor Siiger Prentow, Henrik Blunck, Mikkel Baun Kjærgaard, and Allan Stisen

Department of Computer Science, Aarhus University, Denmark

Abstract. Transportation mode detection is a growing field of research, in which a variety of methods have been developed for detecting transportation modes foremost for outdoor travels. It has been employed in application areas such as public transportation, environmental footprint profiling, and context-aware mobile assistants. For indoor travels the problem of transportation mode detection has received comparatively little attention, even though diverse transportation modes, such as biking, electric vehicles, and scooters, are used indoors, especially in large building complexes.

The potential applications are diverse, and may also extend beyond indoor variants of the above outdoor applications, and include, e.g., scheduling and progress tracking for mobile workers, management of vehicular resources, and navigation support. However, for indoor transportation mode detection, the physical environment as well as the availability and reliability of sensing resources differ drastically from outdoor scenarios. Owing to these differences, many of the methods developed for outdoor transportation mode detection cannot be easily and reliably applied indoors.

In this paper, we explore indoor transportation scenarios to arrive at a conceptual model of indoor transportation modes, and then compare challenges for outdoor and indoor transportation mode detection. In addition, we explore methods for transportation mode detection we deem suitable in indoor settings, and we perform an extensive real-world evaluation of (combinations of) such methods at a large hospital complex. The evaluation presented here utilizes Wi-Fi and accelerometer data collected through smartphones carried by several hospital workers throughout four days of work routines. The results show that the methods can distinguish between six common modes of transportation used by the hospital workers with an F-score of 84.2%.

Key words: Transportation Mode Detection, Indoor Positioning, Mobile Sensing

1 Introduction

Transportation mode detection is a growing field of research, in which a variety of methods have been developed for detecting transportation modes foremost for outdoor travels. It has been employed in application areas such as public transportation, environmental footprint profiling, and context-aware mobile assistants. For indoor travels the problem of transportation mode detection has received comparatively little attention, even though diverse transportation modes, such as biking, electric vehicles, and scooters, are used indoors, especially in large building complexes.

The potential applications are diverse, and may also extend beyond indoor variants of the above outdoor applications, and include, e.g., scheduling and progress tracking for mobile workers, management of vehicular resources, and navigation support. However, for indoor transportation mode detection, both the physical environment as well as the availability and reliability of sensing resources differ drastically from outdoor scenarios. Owing to these differences, many of the methods developed for outdoor transportation mode detection cannot be easily and reliably applied indoors.

In this paper, we explore indoor transportation scenarios to arrive at a conceptual model of indoor transportation modes, and then compare challenges for outdoor and indoor transportation mode detection. In addition, we explore methods for transportation mode detection we deem suitable in indoor settings, and we perform an extensive real-world evaluation of (combinations of) such methods at a large hospital complex. The evaluation presented here utilizes Wi-Fi and accelerometer data collected through smartphones carried by several hospital workers throughout four days of work routines. The results show that the methods can distinguish between six common modes of transportation used by the hospital workers with an F-score of 84.2%.

2 Indoor versus outdoor settings

In this section we discuss the challenges and opportunities provided by respectively indoor and outdoor settings, and how they relate to a representative selection of the many methods for outdoor TMD.

2.1 Transportation Infrastructure

There are significant differences between outdoor and indoor transportation infrastructures, which may influence the results of TMD. Outdoor road networks are often practically unbounded in size, and the possible distance that a tracked person or vehicle may travel is practically unlimited. Indoor travels on the other hand, are bound by a building infrastructure, which greatly limits the sensible travels that may be performed. A result of this difference is that indoor travels are typically much shorter in both distance and time, compared to outdoor travels. For TMD this means that there is less time to determine a specific transportation mode, and that changes in transportation mode will typically happen more often in an indoor setting. In addition the smaller indoor route networks are likely to cause different acceleration profiles through, e.g., more turns and more stops.

Another significant difference between indoor and outdoor transportation infrastructures, is the availability of additional information on the route networks. For outdoor route networks, data is publicly and easily available on roads, biking paths, railways as well as bus stops and routes. This information is very useful for position-based TMD, as it can help limit the number of likely transportation modes [15]. For indoor settings however, this kind of information is much harder to achieve. Floor plans for buildings are only publicly available for a very limited set of buildings, typically

large shopping malls and public places, e.g. railway stations. For other types of buildings, the floor plans may only be available in simple digital formats, such as pictures, which are ill suited for extracting e.g. the route network of hallways in a building [11].

2.2 Transportation Modes

In outdoor scenarios a variety of modes of transportation are common-place; similarly, also indoors, especially in large-scale environments and scenarios, several modes are available some of which overlap with outdoor equivalents. Figure 1 lists what we consider relevant indoor transportation modes. The color of the modes listed indicates, whether they are specific exclusively to indoor (green) or outdoor settings (orange), or whether they occur in variations both in- and outdoors (blue).

Note, that many of the transportation modes (and respectively vehicles) listed could be divided further: e.g. specializations exist for some of them for indoor and outdoor settings, respectively, such as electric versus fossil-fueled forklifts. Indoor transportation modes include various and sometimes specialized electric vehicles which are used as transportation aid in places such as hospitals, airports, or factories, e.g. small electric buses and luggage carts, scooters, bed-pushers and forklifts. These may be designed for and used only indoor, or may be vehicles which can be used in both indoor and outdoor settings.

A significant challenge inherent in most indoor settings relative to outdoor settings is the significantly lower difference in speed profiles of different transportation modes—which is due to foremost the lower top speeds in indoor environments. While it's possible to distinguish between e.g. a bicycle and a car based on maximum speed [15], it's significantly harder to distinguish a bike from an indoor electric vehicle, as the typical speeds fall in similar intervals.

A second challenge in distinguishing indoor transportation modes is that the various modes (resp. vehicles) share the same route network (in contrast to trains versus cars for outdoor travels); additionally most indoor route networks do not cause significant halting patterns during travels such as traffic lights and train stops do.

A further challenge for TMD is the lack of combustion engines in indoor vehicles. Combustion engines in vehicles vibrate at specific frequencies, for example when idling. Distinguishing (fossil-fueled) motorized from other transport modes can thus be facilitated—even when speed profiles for mode candidates are similar—through picking up engine frequencies, which can be achieved even by low-grade, and sub-optimally placed accelerometers, as they are common-place in people's smart phones [23]. Note though that while most indoor vehicle types are not fossil-fueled, recent research reports that electric vehicle motors emit high-frequency sounds which can be picked up by smartphone microphones [19]. This may help in distinguishing electrical vehicles from, e.g., human-powered bikes.

Figure 1 also shows a hierarchical model of transportation modes which we discuss in this paper. Such a hierarchical model gives raise to subdividing the problem of distinguishing (groups of) transportation modes—which brings the advantage of allowing dedicated classifiers for detection decisions on individual nodes, as exploited, e.g. by Hemminki et al. [7] for outdoor TMD. In the hierarchy in Figure 1 higher

entries group similar transportation modes, and consequently neighboring modes in the lower tiers are expected to be more challenging to distinguish.

The root of the hierarchical models resembles the distinction between *stationary* from *moving* activities. For *moving* activities we distinguish two main categories, namely *motorized* and *unmotorized* transport. This distinction is made due to their different qualities with regards to speed and movement patterns, where *motorized* vehicles typically have a different acceleration and speed profile when compared to *unmotorized* vehicles. For *motorized* indoor transports, we consider foremost *electric* vehicles. This category contains general vehicles such as electric *scooters* and *bikes*, but also specialized vehicles for specific settings and applications, e.g., electric forklifts for warehouses or bed-pushers for moving beds around at hospitals. For *unmotorized* transportation, several modes, using e.g. vehicles such as *bicycles* and *scooters*, as well as *walking*, are to be distinguished. The *stationary* activity category is further subdivided into *stationary active*, which covers activities not involving transportation, e.g. performing a stationary work task, while *resting* covers idling, i.e. the absence of significant physical activity.

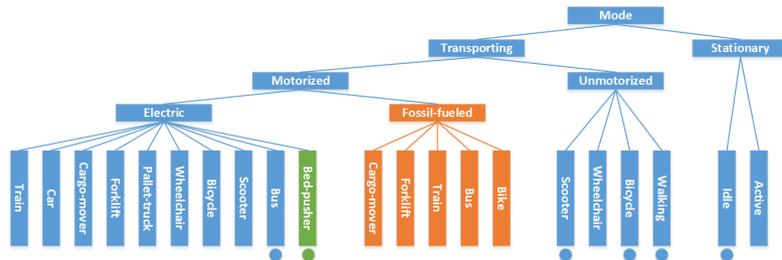
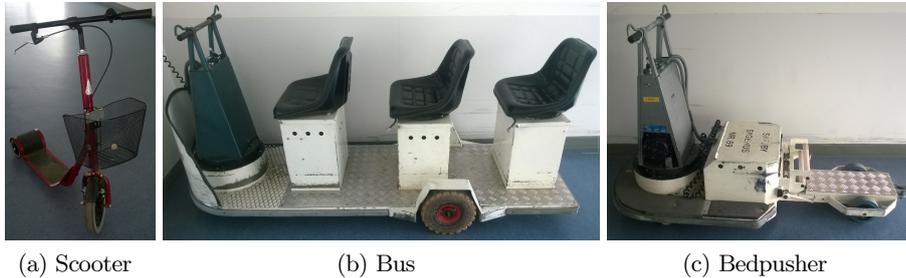


Fig. 1: Hierarchy of transportation modes. Orange: Outdoor-only modes. Green: Indoor-only modes.



(a) Scooter

(b) Bus

(c) Bedpusher

Fig. 2: Common indoor transportation vehicles.

As part of the exploration of indoor transportation modes, we collected transportation mode statistics in a real-world example scenario involving a variety of transportation mode choices, namely the daily work routines of a group of hospital orderlies throughout four days on a large modern hospital complex, covering 150,000 square meter on three floors. The statistics were collected by an assistant assigned to follow the orderlies (without interfering) throughout their work and keep record of the transportation modes they used. In total 300 transportation mode changes were collected, from 6 different modes (as marked in Figure 1): *Stationary*, *walking*, *scooter*, *bike*, *e-bus* and *e-bedpusher*, of which the last two are hospital-specific electric vehicles.

The vehicles can be seen in Figure 2. Table 1 shows the collected statistics. The upper part shows the number of transitions occurring between different transportation modes as a transition matrix, with transitions occurring from rows to columns in the matrix. The lower part lists the total number of trips made with each transportation mode, the total and relative time spent for each transportation mode, as well as the average time for a trip of the given mode.

Transition	Stationary	Walking	Scooter	Bike	E-bedpusher	E-bus
Stationary	0	35	2	1	14	3
Walking	29	0	2	3	64	24
Scooter	2	1	0	0	1	0
Bike	1	4	0	0	0	0
E-bedpusher	16	61	0	0	0	6
E-bus	6	22	0	1	4	0
Total	54	123	4	5	83	33
Time	08:01:30	11:53:21	00:10:32	00:05:01	05:57:27	03:21:55
Percentage	27%	40%	0.6%	0.3%	20%	11%
Average	00:08:27	00:05:47	00:02:38	00:01:00	00:04:15	00:06:07

Table 1: Number of transitions between different modes of transportation.

The table reveals that for all of the transportation modes the average trip duration is rather short (below ten minutes) as compared to typical outdoor scenarios. The mode the persons observed spent most time on during their work routine was walking, with almost 12 hours over four work days. For motorized transportation, the bed pusher is used most often, with about 6 hours, followed by the electric bus, with about 3 hours. The motorized vehicles are clearly preferred compared to the non-motorized bikes and scooters, with in summation only 15 minutes. A significant amount of time is also spend stationary, about 8 hours in total.

Furthermore, the transition matrix reveals that the mode most often transitioned to is walking. This is as expected since walking often serves as an intermediary transportation mode. Direct transitions do occur however, for example when changing between bed pushers and buses. When comparing with results for outdoor settings as presented, e.g., by Zheng et. al.[24] the latter transitions occur less often: Distinguishing the four transportation modes walking, driving, bus, and biking, the authors report about that more than 99% of transitions from modes other than walking occurred towards walking. Indoors the amount of direct transitions between non-walking modes is higher, which is likely due to the fact that outdoors the different transportation vehicles are spatially more separated, e.g., in roads, pavement and biking paths, as well as with regards to parking areas, which means that it is rarely possible to park right next to a bus stop. Indoors the different transportation modes share the same hallways, and to some degree also the same storage places, e.g. electric recharging stations for the electric buses and bed-pushers.

2.3 Available Sensing Technologies

As elaborated above, the sensing capabilities usable for TMD differ significantly between indoor and outdoor scenarios. Table 2 provides a generalised view of the availability and applicability of sensors in indoor and outdoor environments. In the

following, we elaborate on these and their benefits and limitations. Furthermore, we provide pointers into how outdoor TMD make use of sensor modalities, before we overview in Section 3 methods which utilize respective sensor data methods for inferring transportation modes indoors.

Global navigation satellite systems GPS and similar satellite based systems provide—at least in most outdoor environments—ubiquitously available and precise positioning and speed estimates, on which many methods for TMD rely. GPS is not reliably available in indoor settings however, where direct signals from satellites are attenuated, refracted, and reflected by building infrastructure [9]. GPS as a data source is popular within TMD, see .e.g. [24, 12], due to its comparatively reliable and direct speed data coming with accurate position samples in outdoor scenarios [22].

Radio-based sensing Many radio-wave based systems, e.g., WiFi, GSM, Bluetooth and RFID, which are commonly available, e.g. on smartphones, give raise to position estimations. Apart from their use for communication, their signal’s strength can be measured by the participating devices, beacons and mobile clients alike. The resulting *RSSI* (*received signal strength indicator*) data allows for coarse positioning, proximity detection, and for computing further features capturing motion characteristics of mobile clients. The technologies differ in wavelength and emission power, and thus also in range and accuracy. GSM provides very long ranges, up to $34km$, and is almost ubiquitously available, which makes it useful for both indoor and outdoor settings. However, as GSM also has comparatively little spatial variability, the accuracy is typically less than for shorter-range signals, such as Bluetooth or WiFi [10, 21]. With a maximum range of about $50m$ the use of WiFi measurements require a nearby (and preferably dense) network, which makes it most useful for indoor (and urban) settings. Bluetooth and RFID have even lower maximum ranges, which makes them mostly useful for indoor settings or for use with specific gateways, as many beacons, resp. readers, are required to cover a large area. RFID is often used for proximity detection, and can provide very accurate positions, but only for the specific gateway locations where a reader is located. For these technologies the actual RSSI measurements can be performed either by the device or by the network. This makes it possible to track devices with no additional setup on the device, and potentially without the knowledge of the device owner. Maintenance and installation cost is an additional issue that may differ for indoor versus outdoor settings with regards to radio-based infrastructure. For covering outdoor settings, the infrastructure may need to cover a larger area, and beacons to be placed outside need to be resilient to the outdoor environment, potentially at a larger cost.

Kinetic sensors This covers sensors which are capable of sensing motion in different forms, e.g., accelerometers, gyroscopes and magnetometers. As these are commonplace in, e.g., most modern smartphones, they can be used for a variety of application scenarios. They can help to identify movement and acceleration profiles of users to match with transportation modes, as well as be used for dead reckoning in addition to other positioning methods. Specifically for magnetometers however, issues usually arise when using it as a compass indoor, due to significant disturbances in the magnetic

field caused by man-made building infrastructure.¹ Notwithstanding such issues, kinetic sensors are of use in TMD both for outdoor as well as for indoor scenarios, as they do not rely on any environmental sensor infrastructure and do not make strict assumptions about the environment. Thus, the respective TMD methods for outdoor scenarios [5, 7, 12] are useful and comparatively easy to adapt also for indoor scenarios.

Environmental sensors Within this category we group in-device sensors which inform about the user’s (resp. the device’s) current environment, e.g., microphone, light detectors and cameras. These may provide useful information relating to the transportation mode, e.g., microphones can be used to detect specific sounds of vehicles [19], as well as be used for positioning [20], while cameras can be used for positioning and object detection [2]. Such sensors are generally available in smartphones and may be used both indoors and outdoors. However, outdoor settings may provide additional challenges as the environment is less restricted and there are more potential sources of, e.g., noise and light; conversely, distinguishing vehicles outdoors is often easier than indoors due to the noisy engines in, e.g., most cars and buses, c.f. Section 2.2.

	GNSS	Radio-based	Kinetic	Environmental
Outdoor	✓	✓	(✓)	(✓)
Indoor	(-)	✓	✓	✓

Table 2: Sensor types and their applicability for TMD in indoor and outdoor environments.

2.4 Existing Approaches for TMD

This section discusses a representative selection of the many available methods for outdoor TMD, in order to determine their applicability for general indoor settings. Table 3 provides an overview of the selected methods, which are discussed in detail in the following.

Name	Reddy et al. [12]	Sohn et al. [14]	Stenneth et al. [15]	Hemminki et al. [7]
Modes	Stationary, walking, running, biking, motorized	Stationary, walking, driving	Stationary, walking, biking, car, bus, train	Stationary, walking, bus, train, metro, car
Sensors	GPS, accelerometer, (GSM)	GSM	GPS	Accelerometer
External Data	-	-	Bus locations, rail lines, bus stops	-

Table 3: Selected related work on transportation mode detection.

Reddy et al. [12] evaluate the usefulness of several sensors and methods for detecting five different transportation modes. They show that the methods can distinguish between those five modes with an overall accuracy of 93%. However, these results are computed using ten-fold randomized cross-validation, which has a tendency to

¹ In fact, these disturbances are sufficiently significant that they give rise to positioning via fingerprinting instead: given a magnetic field fingerprint collection, a phone’s location can subsequently be estimated within the fingerprinted environment by the local characteristics of the magnetic field as measured by the phone’s magnetometer [1].

produce overly optimistic results, as we will elaborate on in Section 4. The authors conclude furthermore that the best trade-offs between energy-efficiency and accuracy are obtained by choosing GPS and accelerometer as sensors, and that supplying with WiFi and GSM sensor data is not improving accuracy significantly. Their results are instead for employing GPS for speed estimates and the accelerometer for basic statistical and frequency features. The chosen two complement each other well, since using only speed or frequency features does not allow for distinguishing between all transportation modes: E.g., distinguishing stationary mode from motorized transport moving at constant speed may be similar in acceleration patterns, but distinguishable by speed. On the opposite, running and biking can be similar in speed, but have different acceleration variances and dominating frequencies. Their methods are not directly applicable to indoor settings, due to the reliance on speed from GPS. However, one can (and we will) apply the accelerometer features they recommend also in indoor settings.

Hemminki et al. [7] rely solely on the accelerometer sensor for distinguishing still, walking, and four different motorized transport modes. They attribute their improvements in accuracy over related work mostly to a novel method for detecting the direction of gravity, and thereby of the phone’s orientation. Knowing the latter is useful as it eases the comparison of acceleration patterns across phones with different orientations.

Sohn et al. [14] distinguish between three transportation modes: stationary, walking and driving. Their methods are based on collecting consecutive GSM fingerprints, between which the Euclidean distance in signal space is computed. They show that these distances correlate closely with the speed of a device, and achieve an accuracy of up to 90% in distinguishing between the three modes. Their methods take advantage of large expected speed differences of the considered modes—and thus it is not clear how their results generalize to further modes, and especially to indoor settings where most modes are similar in expected speed, c.f. Section 2.2. Their methods, however, are more or less directly applicable to WiFi fingerprints as well.

Stenneth et al. [15] use only GPS measurements, but utilize both speed and position estimates. In combination with external information on real-time bus locations, bus stop locations, and railway layouts, they are able to distinguish the motorized transportation modes into further detail, as either car, bus or train, which they are able to do with an accuracy of up to 93.5%. The external information which they take advantage of is specific to outdoor settings, however it may be possible to employ some similar information from indoor settings, e.g., on locations of electric vehicle charging stations.

3 Features for Indoor Transportation Mode Detection

This section considers the sensing modalities which we consider most useful for indoor transportation mode detection, and presents features which we recommend to extract from the respective sensor data. While we will evaluate the respective features’ usefulness solely in indoor scenarios, note that all of them are also applicable in non-indoor or mixed scenarios. Thus, the listed features can be deployed for outdoor settings without modification, by using the relevant technologies as input, e.g. GPS-retrieved positions instead of positions based on indoor positioning for the position-based features and GSM instead of WiFi for the signal-strength based features. In addition,

further features will be applicable for outdoor settings, such as speed-based features from GPS speed-measurements [12].

The following description of features assumes that the input measurements are aggregated into time windows of a certain duration. We will discuss the size of the windows and other parameter values in Section 4.

3.1 Signal strength based features

As described in Section 2.2, radio infrastructures allows clues about the user’s position, proximity, and motion via measuring incoming signal strengths over time (either on the user device or by the beacons receiving the user device’s radio messages). The signal strength measurements can stem from various sources, e.g., WiFi, Bluetooth or GSM beacons and are structured in *scans*, containing all the received signal strength values from all incoming radio transmitters measured in a single scanning. The frequency of scans, as well as the transmission power, depends on the technology and the specific device settings.

The general idea behind these signal strength based features is that due to the spatial variability of radio signals—and specifically their power loss over distance and when being attenuated by building infrastructure, the signal strength measurements received will vary when moving about—with greater variance when moving fast, e.g., on vehicles. However, variance in signal strength may also stem from other sources such as other people moving in vicinity, opening of doors, etc.

Variance For each window of signal strength measurements, we compute the variance of the signal strength measurements for each beacon. Based on this we use as feature for the window the minimum, maximum and mean values over all beacons.

We compute variance in three forms, differing on the handling of beacons that do not occur in all scans in the window: i) we remove beacons that are not in all scans in the window, ii) we use a low default value for a beacon’s signal strength when the beacon is missing in a scan, and iii) we complete the computation using the values that are available, with no specific handling of missing values.

Access points For each window, we compute the number of beacons from which signal strength measurements have been received. To normalize for different beacon densities in different buildings or different parts of the same building, we also compute the number of beacons divided by respectively the average and total number of received signal strength measurements during each scan. The motivation behind these features is that the number of beacons from which measurements have been received serves as a measure of the size of the area covered in the time window at hand. In addition, we compute the number of changes in beacons, as the total number of either appearing or disappearing beacons in consecutive scans in the window. Given the fixed duration of the time window at hand, the above features serve as a proxy for speed.

Distance in signal space For each pair of consecutive scans in each window, we compute the Euclidean distance in signal space between the samples. This is done by treating each access point as a dimension in the Euclidean distance formula, as described by Bahl et al. [3]. As features we extract the mean, minimum and maximum of the distances computed within a window.

3.2 Position based features

When a positioning system is available, we can use the (sequence of) timestamped position estimates to extract features. The usefulness of these features obviously depends on the accuracy of the positioning system used.

Movement speed We can compute an estimate of the current speed from (consecutive) position estimates. To mitigate the effects of positioning inaccuracies, we smooth the position sequence by computing the median trajectory of the given positions [4]. We then use these smoothed positions to compute the speed between each consecutive position estimates, from which we for each window compute the mean, maximum, minimum and end-to-end speed in the window, which are used as features.

3.3 Kinetic features

Time domain The kinetic sensors we used include time-domain features which have been used for human activity recognition (HAR) in multiple systems, as described by Figo et al. [5]. The time-domain features include statistical measures such as mean, standard deviation, minimum and maximum values and root mean square. The full list of features used here is given by Figo et al. [5, Table 2].

For the time domain features, as for the following two classes of kinetic features, we employ accelerometer measurements from each axis x, y, z as well as the orientation-independent magnitude $\sqrt{x^2 + y^2 + z^2}$.

Frequency domain The kinetic features subsumed here are based on frequency analysis, which is useful for detecting (especially periodic) patterns of movement, such as walking. These features include the normalized spectral coefficients from 0 to 20 Hz, the entropy of coefficients and the dominating frequency, see Figo et al. for a full list of features used here [5, Table 3].

ECDF features Additional to the explicitly chosen kinetic features above, we also employ feature learning in the form of *empirical cumulative distribution function (ECDF)* features, as described by Hammerla et al. [6] for accelerometer measurements. Note that while the kinetic features above can also be employed for kinetic sensors other than accelerometers, such as gyroscopes or magnetometers, ECDF features are expected to generalize even better—as they adapt to the patterns in the provided training data.

4 Evaluation

This section details the evaluations we have performed of the methods for indoor TMD described in section 3. We evaluate the methods based on real-world data collected from a large hospital, as described in section 4.2.

For evaluating approaches for indoor TMD, we undertook a real-world evaluation in a concrete case setting where accurate TMD provides significant value. Advantages of this choice includes the evaluation in the face of both real-world challenges, varying use-case scenarios and user behaviours, as well as a discussion of the real-world impacts of shortcomings in TMD accuracy and reliability.

4.1 Evaluation setting

The scenario in which we evaluate indoor TMD methods lies within mHealth, specifically within hospital task logistics and focuses on the automatic scheduling and registering of tasks of orderlies. To this end, various prototypes have been evaluated in cooperation with hospital stakeholders and staff at several hospitals, including the one chosen as environment for this investigation [17]. To further improve the support of hospital work logistics, the automatic detection of which transportation modes the orderlies use promises the following improvements for the overall hospital logistics:

- *Travel time estimation:* Inferring transportation modes from collected data in real-time will provide more accurate travel time estimates for the current as well as for potential future tasks. That will in turn render task scheduling more efficient.
- *Navigation:* The available routes through a building are restricted by the transportation mode used, e.g., stairs cannot be used when on a vehicle. Thus, for navigational aids knowing the user’s current mode is essential.
- *Task registration:* For automatic registration of tasks, detecting changes in transportation mode provides valuable clues for automatic trip detection—and may thus serve to detect the begin or ending of a task, or a task phase, or for determining which (sequence of) task types are being performed.
- *Vehicle management:* For managing upkeep for a fleet of vehicles, the amount, rate, and areas of use that each individual vehicle and each type of vehicle has seen can provide valuable input, both for scheduling maintenance as well as for optimizing the size of and default locations for the vehicle fleet.

4.2 Dataset

The dataset used comprises the activities of several orderlies throughout their daily work routines for four days—which include maintenance work, transporting patients, and similar tasks, which may require different transportation means, c.f. Section 2.2. When collecting data with an orderly, we supplied him with a smartphone which logged accelerometer and WiFi-scan signal-strength measurements, at the highest available frequencies. This was approximately $0.5Hz$ for WiFi, and $200Hz$ for accelerometer. The smartphone was carried in a shirt pocket where the orderlies would normally carry their phones, so we do not evaluate for different phone placement locations.²

Meanwhile, ground truth with regards to used transportation means was collected by a person following the orderly throughout his work. The person was following on an independent vehicle such as a bike, so as to influence neither the orderlies (and specifically: their choice of transportation means), nor the collected sensor measurements. We also had access to the WiFi network of the hospital, which allowed us to collect all signal-strength measurements made by the network for signals from the phone carried by the orderly.

² While the assumptions of homogeneity in device placement and smartphone model are valid in the use scenario of this study, such homogeneity may be missing in other scenarios and lead to lower accuracies for distinguishing transportation modes, see, e.g., [16, 7].

4.3 Evaluation methodology

For the evaluation, in line with standard practices in activity recognition research, we use the F_1 -score, which is the harmonic mean of precision and recall [13].

The evaluations are performed using 10-fold cross-validation. However, we do not use random folds for the cross-validation, as random folding is prone to provide overly optimistic results when dealing with time-series data. Due to being collected close in time, neighbouring measurements will often be similar, but may be selected for different folds. Thus the classifier will be trained on measurements that are similar to those which are tested on. Instead we split folds on time, so that the first fold will be the first tenth of the data time-wise, etc. This ensures that the folds are completely independent, as they will stem from different trips made by the orderlies. For comparison we performed the evaluation of Section 4.4 also using random-fold cross-validation which resulted in an unrealistically high F-score of 99.2%, as compared to the F-score of 84.2% when using time-folded cross-validation.

We evaluate four different classifiers that are popular for use in human activity recognition, namely nearest neighbour classifier with $k=5$, C4.5 decision tree, support vector machines (SVM), and random forest.

4.4 Detecting transportation means

In this section, we present overall accuracies for detecting different transportation modes, in different layers of the transportation mode model, as shown in Figure 3. Two hospital-specific vehicles are included: The e-bus and the e-bedpusher. Both are electric vehicles which serve two different purposes. The e-bus is used for fast transport by an orderly, as well as up to two passengers which are typically patients at the hospital. The bed-pusher is designed for use by an orderly for fast and easy transportation of bed-bound patients - however the orderly may also use it for transporting just himself.

For the evaluation we used a random forest learner, parametrized to train 100 trees, and a window size of 10s for accelerometer measurements, and for Wi-Fi measurements we use a hierarchy of window sizes: a short one of 10s and a large one of 60s. Later on we will discuss the influence of the chosen parameters, specifically for various classifiers in Section 4.6, and for various window sizes in Section 4.8. Complementing Figure 3, Figure 4 shows the resulting confusion matrix. Figure 3 shows the accuracy in the different layers of the transportation mode model. Note that F-scores given at each node refer to the task of distinguishing that node’s mode from just its siblings’ modes. It is visible, that when only distinguishing between categories of upper-level layers, distinguishing is done accurately, e.g., for the top-level categories “transporting” and “stationary”—with an F-score of 98.6%. Considering more specific transporting subclasses, for distinguishing between motorized and unmotorized transportation, the accuracy falls to an F-score of 92.8%. This fall continues when distinguishing between the two quite similar motorized vehicles considered: bus and bedpusher—with an F-score of only 61.0%. For distinguishing unmotorized transportation, results are better: walking is detected correctly 98.7% of the time, whereas the scooter and bike transportation - for which we have only little data - are harder to detect. This may also be

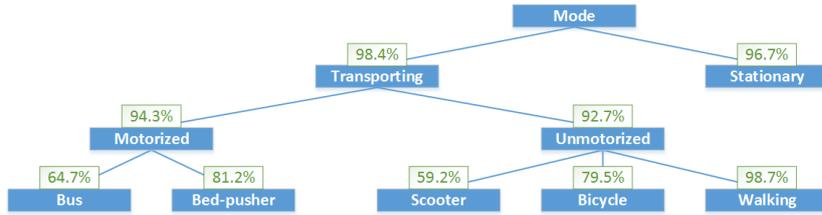


Fig. 3: Accuracy in detecting different levels of transportation means.

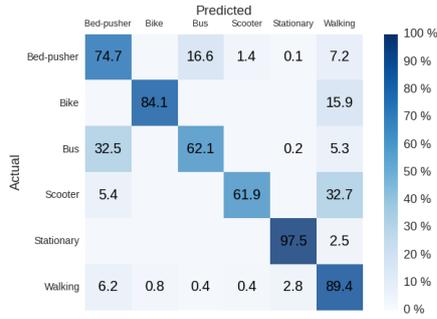


Fig. 4: Confusion matrix based on accelerometer and WiFi features

due to that the movements may look similar to walking, e.g. for powering a scooter. In total the modes of the unmotorized category can be detected with an F-score of 97.5%.

To a large degree the results match the expectations. The features can accurately determine whether the user is moving or not, since the typical values of both WiFi and kinetic features are well separated for these two cases. The motorized and unmotorized categories can also be distinguished quite clearly, likely due to the fact that a person riding a motorized vehicle will be standing still on the vehicle, while a person riding an unmotorized vehicle will move in order to power the vehicle, or in order to walk - causing characteristic (and usually: periodic) movement patterns. The distinction between bus and bed-pusher, however, was likely to be hard—as both vehicles are quite similar, specifically they are both electrically-powered and drive at about the same speeds.

4.5 Evaluating different sensor and feature types

In this section, we evaluate the usefulness of the different feature and sensor types, as well as of individual features, for indoor TMD.

Comparing the usefulness of different feature types We will do so by evaluating—for all the transportation modes for which we have collected data—various candidates using different combinations of feature types. Figure 5a shows the resulting confusion matrix when using only WiFi features, while Figures 5b, 5c and 5d show confusion matrices for employing solely accelerometer features of one of the three kinetic feature types, respectively: ECDF, time-domain and frequency-domain features. Table 4 shows F-scores for use of individual feature-types as well as for any combination of features, ordered by F-scores.

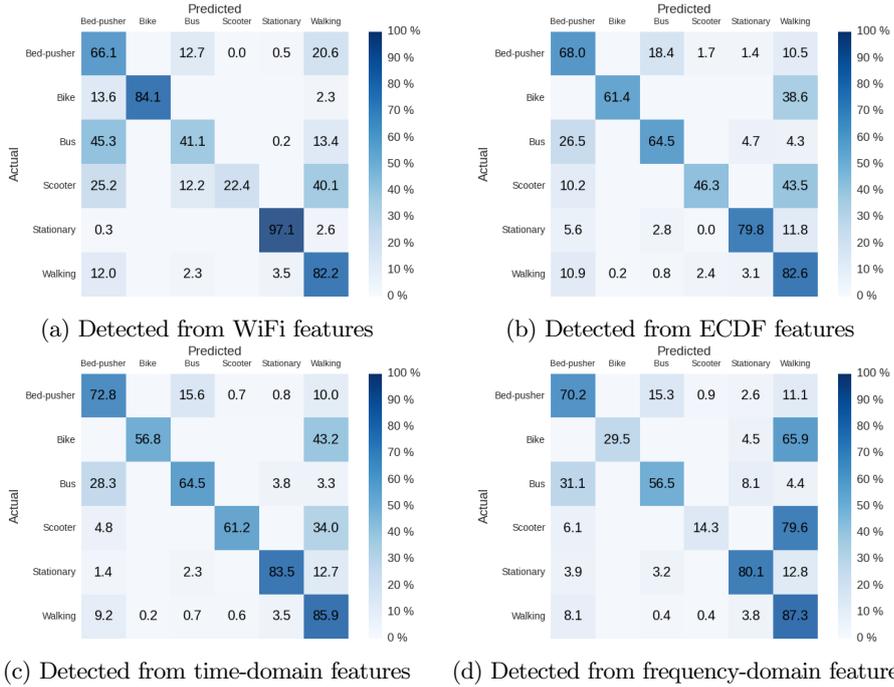


Fig. 5: Confusion matrices holding F-scores when using a single feature type.

ECDF	-	-	-	✓	✓	✓	✓	-	-	✓	✓	✓	-	-	✓
Frequency	-	✓	✓	✓	-	-	✓	-	✓	-	✓	✓	-	✓	-
Time	✓	✓	-	-	✓	-	✓	✓	✓	✓	✓	-	-	-	-
WiFi	✓	✓	✓	✓	✓	✓	✓	-	-	-	-	-	✓	-	-
Result	84.2	83.9	83.5	83.1	83.0	82.8	82.8	78.9	78.9	78.6	77.7	76.0	75.9	75.7	

Table 4: The results for different combinations of feature types

For individual feature types the results show that the type providing the best results for classification of these indoor transportation modes are the kinetic time-domain features with an F-score of 78.9%, the next best being WiFi features with an F-score of 76.0%, followed by ECDF and frequency-domain features. These results are mirrored in the combined features, where the best results are achieved by a combination of time-domain accelerometer features and the WiFi features, with an F-score of 84.2%. Interesting to note is that for all combinations of accelerometer-based features, the inclusion of WiFi features improves the results.

The individual figures show that accelerometer has better performance than WiFi when distinguishing walking from the motorized vehicles. As the vehicles may at times drive at speeds close to walking, it may be hard to distinguish them based on WiFi features, while the accelerometer features will be much more distinct due to the specific movement patterns from walking. On the other hand, WiFi features are significantly better for distinguishing stationary from walking, as they can distinguish shifting weight between legs or shuffling around from walking a distance. In this way the ability of accelerometers to detect very small-scale movements complement the ability of WiFi to detect major movements.

Comparing the importance of individual features In order to evaluate which particular features appear most useful, we measured accuracy changes resulting from including (vs. excluding) individual features. Figure 6 shows for each of the features types—WiFi, ECDF, time-domain, and frequency-domain features—the five features which by themselves cause the highest accuracy gains. The most important ones among the WiFi features we used, are based on the number of unique access points seen in a time window—which is a rather crude proxy of the area covered in that window, and in turn for speed. Less well performing are WiFi features based on variance in signal strength measurements, or on the estimation of speed via WiFi positioning data (likely due to the inaccuracy of the latter). For the ECDF features notably the higher bins appear important, suggesting that the values in the high end of the histogram are best for distinguishing transportation modes. From the time domain kinetic features, variance and standard deviation, as well as the 75 percentile of the normalized measurements show high importance. As with the well performing ECDF features, the feature values’ independence of phone orientation is likely to be helpful. Although, also useful appear the variance and standard deviation of measurements on the y -axis—which is for normal phone orientation the axis which captures upwards and downwards movement best, and is thus helpful for detecting, e.g., steps during walking. For the frequency domain features, features covering movements in the frequency range from 0 to 2 Hz are dominating, which is a bandwidth which captures well the characteristics of many common human movement patterns [18].

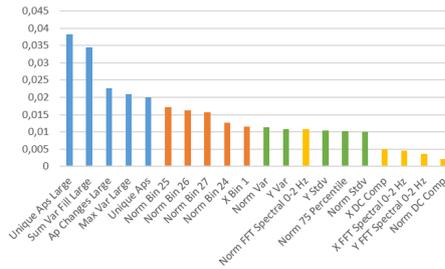


Fig. 6: Computed feature importance for the 5 most important features of each type. Blue: WiFi, Orange: Accelerometer-ECDF, Green: Acc-Time, Yellow: Acc-Frequency

4.6 Evaluating different classifiers

We evaluate several candidate classifiers in regards to their usefulness for indoor TMD when using combined accelerometer and WiFi features. We selected candidates that are often used in similar classification tasks such as activity detection and TMD, as described in Section 3. For each candidate, we have experimented with the relevant parameters; we obtained best results, e.g. for $C4.5$ tree classifier when setting a max-depth of 10, and for the random forest, results did not improve when using more than 150 trees. For the K -nearest neighbours algorithm, we found $K = 3$ to work best. Table 5 presents F-scores of time-folded cross-validation for each of the classifiers. We see that random forests seem to be superior to support vector machines and K -nearest neighbour classifiers in this setup, while all are superior to the $C4.5$ classifiers. A likely reason is that the setup is prone to overfitting, and that thus the

more resilient classifiers perform better; e.g. random forests profit from the several trained trees voting to reach a common verdict. These results and explanations are in line with those reported by Sagha et al. [13] for indoor activity recognition.

Learner	C4.5	K-NN	SVM	Random forest
F_1 -score (%)	78.6	81.2	81.6	84.1

Table 5: F-scores when using different classifiers.

4.7 Network- versus client-based Wi-Fi signal strength measurements

Access to the WiFi network infrastructure at the hospital site allowed us to evaluate also (the differences between) using network-based and client-based WiFi signal strength data collection. The results showed an only marginal F-score decrease of 0.2 when using network-collected instead of on-phone-collected measurements—and thus, that accurate TMD is also possible when using network-based WiFi, i.e., without requiring the users’ mobile devices to perform measurements. Note that in the network-based setup evaluated here we do however depend on the mobile device to frequently send out WiFi signals, through WiFi scans or other data communication.

4.8 Evaluating different Wi-Fi window size

Table 6 presents the F-scores when varying the time window size over which we aggregate signal strength data to obtain WiFi feature values from 10 up to 640 seconds. The accuracy improves significantly when increasing a smaller window size, but the increase lessens once sizes reach a minute, and using a window size of 10 minutes provides for worse results than for 5 minutes. The optimal choice for a window size relates to the expected trip length. As shown in Table 1, the trips performed are on average below 9 minutes for all transportation modes. Increasing the window size beyond yields high chances that several transportation modes are covered in a given time window. Additionally, increased window sizes also increase the response-time for detecting transportation mode changes in a real-time setting.

Window size	10s	20s	40s	1m20s	2m40s	5m20s	10m40s
F_1 -score (%)	66.8	69.1	72.6	77.1	78.1	81.4	80.3

Table 6: F-scores for the different window sizes of WiFi features.

4.9 Evaluating the usefulness of additional tagging of vehicles

For cases where the transportation modes are hard to distinguish due to similar movement profiles, e.g., for bus versus bed-pusher vehicles, the accuracy can be improved by fitting WiFi tags onto such vehicles. To evaluate the usefulness of WiFi tags, we fitted a bus and a bed-pusher with tags that performed WiFi scans every 10 seconds. These scans were collected by the WiFi infrastructure during our observation of the orderlies. From the collected scans we computed as features the Euclidean distance in signal space of the orderlies’ WiFi measurements to those of respectively the bus and the bed-pusher WiFi tags—the intuition here being that when the order drives, e.g. a bus, he will be in close proximity to the bus tag, resulting in a low signal space distance. Figure 7 shows an example of the computed signal space distance

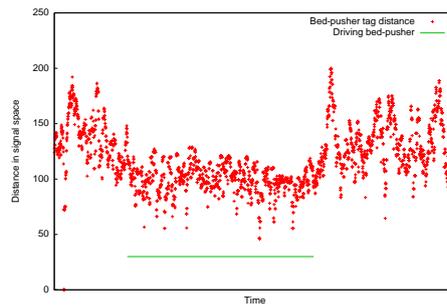


Fig. 7: Distance in signal space between the user and tag placed on bed-pusher.

before, during, and after a trip on a bed-pusher. It shows how the distance decreases when the user approaches the tag, and it is consistently low while he drives the bed-pusher. Our evaluation showed that with the addition of tags, the distinction between bus and bedpusher improved from an F-score of 49.3% to an F-score of 82.3%. However, this approach may not work in situations where several types of vehicles are driven in closer proximity—in which case co-movement detection could be used on sequences of position estimates for further disambiguation [8].

5 Conclusions

Through the performed evaluations we have shown that automatic detection of transportation modes using mobile devices such as smartphones is possible also in indoor settings. The results show that accelerometer or WiFi measurements can be used individually to detect transportation modes, but that a combination of both provides for improved results: Accelerometer-based features show strength in distinguishing between walking and vehicles, while WiFi is especially useful for determining whether a user is moving or not, and for distinguishing between vehicles used in case those are WiFi-tagged. Future work includes automatic detection of transportation modes also in mixed indoor and outdoor setting, in order to support e.g., logistics in such settings. Furthermore, also additional sensors and respective fusion techniques, e.g. orientation-aware motion sensor features, may provide further accuracy gains. Finally, it is left as future work to explore the impact of different device placements and models on the classification accuracy.

References

1. Indoor Atlas. <http://www.indooratlas.com>. Accessed: 2015-08-03.
2. D. C. Asmar, J. S. Zelek, and S. M. Abdallah. Smartslam: localization and mapping across multi-environments. In *Proc. Intl. Conf. Systems, Man and Cybernetics*, 2004.
3. P. Bahl and V. N. Padmanabhan. Radar: An in-building rf-based user location and tracking system. In *Proc. IEEE Conf. Computer Communications*, pages 775–784, 2000.
4. K. Buchin, M. Buchin, M. van Kreveld, M. Lffler, R. Silveira, C. Wenk, and L. Wiratma. Median trajectories. *Algorithmica*, 66(3):595–614, 2013.

5. D. Figo, P. C. Diniz, D. R. Ferreira, and J. M. P. Cardoso. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*, 14(7):645–662, 2010.
6. N. Y. Hammerla, R. Kirkham, P. Andras, and T. Plötz. On Preserving Statistical Characteristics of Accelerometry Data using their Empirical Cumulative Distribution. In *Proc. ISWC '13*, 2013.
7. S. Hemminki, P. Nurmi, and S. Tarkoma. Accelerometer-based transportation mode detection on smartphones. In *ACM SenSys 2013*, pages 13:1–13:14. ACM, 2013.
8. M. B. Kjærgaard and H. Blunck. Tool support for detection and analysis of following and leadership behavior of pedestrians from mobile sensing data. *Pervasive and Mobile Computing*, 10:104–117, 2014.
9. M. B. Kjærgaard, H. Blunck, T. Godsk, T. Toftkjær, D. L. Christensen, and K. Grønbæk. Indoor positioning using GPS revisited. In *Pervasive 2010*, pages 38–56, 2010.
10. A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello, and B. Schilit. Place lab: Device positioning using radio beacons in the wild. In *Pervasive Computing*, volume 3468 of *Lecture Notes in Computer Science*, pages 116–133. 2005.
11. T. S. Prentow, A. Thom, H. Blunck, and J. Vahrenhold. Making sense of trajectory data in indoor spaces. In *IEEE 16th Intl. Conf. Mobile Data Management*, 2015.
12. S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava. Using mobile phones to determine transportation modes. *ACM Trans. Sen. Netw.*, 6(2):13:1–13:27, 2010.
13. H. Sagha, S. Digumarti, J. del R Millan, R. Chavarriaga, A. Calatroni, D. Roggen, and G. Tröster. Benchmarking classification techniques using the Opportunity human activity dataset. In *IEEE Systems, Man, and Cybernetics (SMC)*, 2011.
14. T. Sohn, A. Varshavsky, A. LaMarca, M. Chen, T. Choudhury, I. Smith, S. Consolvo, J. Hightower, W. Griswold, and E. de Lara. Mobility detection using everyday gsm traces. In *UbiComp 2006*, volume 4206, pages 212–224. 2006.
15. L. Stenneth, O. Wolfson, P. S. Yu, and B. Xu. Transportation mode detection using mobile phones and gis information. In *Proc. 19th ACM GIS*, pages 54–63. ACM, 2011.
16. A. Stisen, H. Blunck, S. Bhattacharya, T. S. Prentow, M. B. Kjærgaard, A. Dey, T. Sonne, and M. M. Jensen. Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition. In *ACM SenSys 2015*. ACM, 2015.
17. A. Stisen, N. Verdezoto, H. Blunck, M. B. Kjærgaard, and K. Grønbæk. Accounting for the invisible work of hospital orderlies: Designing for local and global coordination. In *ACM CSCW 2016*. ACM, 2016.
18. M. Sun and J. Hill. A method for measuring mechanical work and work efficiency during human activities. *Journal of biomechanics*, 26(3):229–241, 1993.
19. M. Takagi, K. Fujimoto, Y. Kawahara, and T. Asami. Detecting hybrid and electric vehicles using a smartphone. In *ACM UbiComp '14*, pages 267–275, 2014.
20. S. P. Tarzia, P. A. Dinda, R. P. Dick, and G. Memik. Indoor localization without infrastructure using the acoustic background spectrum. In *Proc. MobiSys 2011*, 2011.
21. A. Varshavsky, E. de Lara, J. Hightower, A. LaMarca, and V. Otsason. GSM indoor localization. *Pervasive and Mobile Computing*, 3(6):698–720, 2007.
22. T. Witte and A. Wilson. Accuracy of non-differential GPS for the determination of speed over ground. *Journal of Biomechanics*, 37(12):1891 – 1898, 2004.
23. M. Wüstenberg, H. Blunck, K. Grønbæk, and M. B. Kjærgaard. Distinguishing electric vehicles from fossil-fueled vehicles with mobile sensing. In *IEEE MDM 2014*.
24. Y. Zheng, Y. Chen, Q. Li, X. Xie, and W. Ma. Understanding transportation modes based on GPS data for web applications. *TWEB*, 4(1), 2010.