A Study on the Impact of Indoor Positioning Performance on Activity Recognition Applications

Andreas Mathisen
Aarhus University
Denmark
am@cs.au.dk

Søren Krogh Sørensen
The Alexandra Institute
soren.krogh@alexandra.dk

Allan Stisen
Aarhus University
Denmark
allans@cs.au.dk

Henrik Blunck
Bochum
University of Applied Sciences
Germany
henrik.blunck@hs-bochum.de

Kaj Grønbæk
Aarhus University
Denmark
kgronbak@cs.au.dk

ABSTRACT

Due to substantial research in indoor positioning a vast amount of location technologies and algorithms are available to enable various applications. However, it is challenging knowing which positioning system is optimal, or sufficient, for a specific application—not only when considering development and maintenance costs, but also the potential impact of the positioning system's performance on the respective application's performance.

In this paper, we present an evaluation of how positioning performance in various system setups affects two chosen real-world applications at a $160,000m^2$ hospital building complex using an existing WiFi system as primary sensing infrastructure. While a multitude of positioning applications are run at the hospital, we focused on two applications within human activity recognition (HAR). HAR is an application area of positioning where the impacts are indirect and challenging to predict, therefore motivating this type of investigation yet underrepresented in the literature.

Our evaluation includes several WiFi based indoor positioning system variants, with different accuracies, and investigates the impact of their respective performances on the two HAR applications. Among others, our evaluation shows that positioning accuracy, as measured traditionally, is not the only performance measure important to consider, and that it has a surprisingly small impact on the applications' performance—suggesting that the increased costs for a higher-accuracy positioning system often do not yield the anticipated returns.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing design and evaluation methods**; **Empirical studies in ubiquitous and mobile computing**;

MobiQuitous 2017, November 7–10, 2017, Melbourne, VIC, Australia

ACM ISBN 978-1-4503-5368-7/17/11…$15.00
https://doi.org/10.1145/3144457.3144501

ACM Reference Format:
Andreas Mathisen, Søren Krogh Sørensen, Allan Stisen, Henrik Blunck, and Kaj Grønbæk. 2017. A Study on the Impact of Indoor Positioning Performance on Activity Recognition Applications. In *MobiQuitous 2017: the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, November 7–10, 2017, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3144457.3144501

1 INTRODUCTION

Indoor positioning technologies have enabled a widespread use of location sensing, and indoor positioning is often a primary resource for applications in various domains, where, e.g., higher-level sensing-based recognition tasks such as Human Activity Recognition (HAR), or in-building occupancy detection, is employed. For many of these applications an important concern is the accuracy provided by the utilized positioning technology and for many of these systems there is an underlying hypothesis that an increased positioning accuracy will result in more accurate applications. However, higher-accuracy positioning technologies often come with higher costs for deploying and maintaining the respective indoor positioning systems. These costs can include deploying new sensing technology, time spent on designing, developing and evaluating algorithms, or gathering more training data as needed for empirical fingerprinting, etc.

As such, it is vital for applications that an adequate indoor positioning technology is chosen, rather than an overly costly one. However, to the best of our knowledge, no studies have yet rigorously investigated the influence of positioning performance on HAR applications. To bridge the gap between indoor positioning as a standalone problem and the assessment of its usefulness to other applications, we need to understand how positioning performance affects specific use scenarios. In this paper, we take a step towards such an understanding by evaluating several positioning systems—all of them based on WiFi-infrastructure but yielding different positioning performance characteristics—in regards to their usefulness for HAR applications. We investigate two specific HAR applications, namely Task Phase Recognition (TPR), which is the recognition of phases within tasks executed by mobile workers, and Indoor Transportation Mode Detection (ITMD), i.e. differentiating movement by, e.g., foot, bike, scooter and various e-vehicles. In earlier work [20, 27], we reported on solutions for these problems. These solutions were deployed and evaluated in several hospitals, and we describe them in more detail in Section 4 as well as how they are used in logistical operations and for mobile work-flow awareness.

While there is a wide array of applications of indoor positioning systems, and even the aforementioned hospitals run a multitude of them in parallel to TPR and ITMD, we chose to focus on the latter two applications for three reasons. Firstly, HAR applications allow for clearly quantifiable assessment measures, thus allowing also to assess the impacts of positioning performance. Secondly, we argue that TPR and ITMD are representative of HAR applications which strongly rely on positioning as input, and thirdly, our experiences and investigations have shown that the aforementioned impacts and correlations between positioning and HAR performance are complex and thus worthwhile investigating.

Concretely, at the hospital test-bed, both TPR and ITMD rely for contextual input (among other data types) on position tracking data. The two applications make no explicit requirements on the positioning system used and currently input positions are inferred from Received Signal Strength Indicators (RSSI) measurements from WiFi; additionally also the raw RSSI measurements are used a contextual input. Noteworthy, the TPR and ITMD problems differ in the relative importance of these two inputs: In the case of TPR, indoor positioning data was identified as most beneficial for high recognition accuracy, while for ITMD raw RSSI data proved more beneficial, since the position estimates obtained were deemed too inaccurate as motion speed indicator compared to other proxies for speed. For both the TPR and the ITMD tasks, we evaluate comparatively the respective effects of using a variety of WiFi positioning systems, ranging in median accuracy from 10.95 to 6.51 meters.

Overall, the main goal of this paper is to enhance the understanding within the research areas of both indoor positioning and HAR, by investigating empirically, using TPR and ITMD as example HAR applications, how positioning performance influences recognition accuracies. Interesting observations we report on include that when switching from a higher- to a lower-accuracy positioning solution, for both TPR and ITMD we saw no or only small reductions (of only a few percentage points in F_1-score) in recognition accuracy. Additionally, our results indicate that not only positioning accuracy impacts the recognition accuracy, but that other performance metrics, such as coherence in speed and travel bearing between position estimates are also or even more indicative of resulting HAR performance.

2 RELATED WORK

Indoor positioning has received increasing attention from the research community over the last decade. However, Kjærgaard et al. [11] point out that a widespread breakthrough of successful real world deployments has not yet been achieved. Their findings suggest that the narrow research focus on algorithmic optimizing of positioning accuracy for small-scale test beds is insufficient in realistic settings where organizational requirements, like low maintenance effort, play a large role. The insights presented in [11] are also reflected in evaluations of positioning systems—which traditionally are limited to small office buildings and few evaluation metrics, whereas recently larger-scale evaluations are called for [19, 30].

A multitude of different approaches have been proposed for indoor positioning, ranging from low effort methods, such as basic lateration [21], to high-effort methods, like empirical fingerprinting combined with inertial sensors [4]. Building on existing infrastructure is a simple way to lower deployment efforts, which explains the large focus on indoor positioning using received signal strength (RSS) from WiFi. A main challenge inherent in these methods is to accurately model signal attenuation. Bahl et al. [1] propose, not only emperical fingerprinting, but also as an alternative a purely model-based approach, which has achieved wide acceptance as it does not require manual fingerprint collection. Combining model-based approaches [1] and empirical fingerprinting [9, 10], some approaches propose to learn signal attenuation in a specific environment based on smaller empirical data sets [7].

Similar approaches as for WiFi have been proposed also for other technologies: Ni et al. [16] uses RFID tags in a dense grid. Other signal-based approaches employ IR [28], ultrasound [22], foremost for room level positioning, and more recently bluetooth low energy infrastructures [5]. While the algorithmic variants (and thus also the focus of this investigation) is applicable to most of the above non-WiFi approaches, the the relatively large deployment cost for the latter limits the use scenarios where such systems are feasible.

Generally, there is yet little research into how the mentioned positioning approaches will perform in various use-scenarios, especially when these involves further data processing trough other components such as HAR systems—which is an important consideration in order to minimize the gap between research and real world deployments of indoor positioning systems. HAR has been studied for various work domains and has been proposed with many different foci. Kwapisz et al. [13] propose activity recognition in the classic sense, i.e. to detect human activities, such as walking and jogging, using inertial sensors from mobile phones. Ward et al. [29] present activity recognition of assembly tasks using microphones in addition to inertial sensors. Stiefmeier et al. [25] study wearable activity tracking of assembly line workers in car manufacturing and Koskimaki et al. [12] also propose activity recognition for industrial assembly lines based on wearable sensors, i.e. a wrist-worn inertial measurement unit. Activity recognition has, furthermore, been proposed for the clinical setting and identified as key to realize the smart hospital [24]. Favela et al. [6] estimate high-level activities in medical wards using manually labeled contextual information from an observational study. Parlak et al. [18] presents a system that can recognize activities during trauma resuscitation using RFID sensors and tagged medical items. Recognizing surgical steps and activities have also been studied. Padoy et al. [17] propose workflow monitoring using 3D motion features, where they are able to recognize phases using image analysis. Bardram et al. [2] are also able to predict phases of an operation based on body-worn sensors, e.g. RFID readers and tags.

While activity recognition has been proposed for many domains and with many different types of input data, there have been few studies on how variations in the contextual input data influence the recognition accuracy in different systems.

3 INDOOR POSITIONING METHODS CONSIDERED

In this paper, to assess the impact on HAR of real-world positioning performances, we investigate a range of positioning methods and filters, as we deem suitable for large-scale deployment scenarios.

The range investigated additionally includes artificial degradation with various intensities of the estimated position estimates.

In such scenarios it is often crucial that development and maintenance efforts of new software systems are kept low, as it is often simply infeasible to maintain, e.g., an empirical fingerprint map or to deploy and maintain large quantities of additional hardware [11]. With the following methods, we obtain a wide range of positioning performance levels and characteristics such as accuracy, precision and temporal coherence, to increase the real-world applicability of our evaluations.

We only briefly describe the positioning methods investigated here, for a complete description we refer to earlier work, describing the methods as well as experiences and evaluations as undertaken at one of the deployment sites, namely a large hospital complex in Denmark also used as test bed for this artcle's investigation [14].

Weighted Centroid (WC): position estimates are computed as the average position of the k strongest WiFi access points using the received signal strength indicator as weight.

Radio Map (RM): position estimates are computed with model based fingerprinting, where coarse images of the indoor environment has been used to estimate the radio map.

Temporal Filter (TF): position estimates are smoothed by pulling an outlier towards the previous position estimate, where an estimate is deemed an outlier if the speed needed to reach this estimate exceeds a threshold.

Particle Filter (PF): position estimates are smoothed by a particle filter using coarse images of the indoor environment and with random movements, i.e. no inertial sensors is utilized to infer direction or speed.

At the test site hospital complex, several environmental difficulties were identified emphasizing the existing gap between reported accuracies in research and what can be expected in the real world. Furthermore, the work presented in [14] already highlights that positioning performance can not be described by a single evaluation metric: While positioning systems are often compared just by their positioning accuracy, other performance measures such as coverage and trace coherence in terms of speed or angular changes are important. Table 1 lists for the six positioning system variants deployed at the test-bed hospital their mean and median accuracies, as previously reported [14], as well as two measures for (in-)coherence in speed and bearing. To compute these latter, estimate trajectories and ground truth trajectories are aligned wr.t. time, i.e. are interpolated at 1Hz, which allows to match one-second segments of estimated vs. ground-truth trajectory. The speed coherence measure is computed as the average difference in one-second segment length (in meters) between the estimate trajectory and the ground truth trajectory, i.e. the speed coherence measure describes the average difference in m/s when traveling along the two trajectories. The angle coherence measure is calculated as the average difference in bearing (in degrees) for the segments of the two trajectories.

Noteworthy, it is not always the case that the most accurate system also outperforms the competitors in the coherence measures. As expected, filtered positioning methods show significantly higher coherence in Table 1, and, also as expected, more so for more elaborate (here: particle) filters. Note also, that WC shows higher coherence values than RM; this is expected as WC inherently is defined by averaging observations, i.e. RSSI values, which can be expected to develop rather coherently over time. In contrast, RM-fingerprinting, identifying the RSSI fingerprints best resembling the current RSSI measurements, may potentially induce incoherent "jumps" in reported positions (in case the best resembling fingerprints are found far apart for consecutive RSSI measurements). When artificially degrading positioning performance, similar observations apply, and we will present results for degradations of different levels in Section 6.3.

Performance measure	WC	TF(WC)	PF(WC)	RM	TF(RM)	PF(RM)
Acc. (E_u) Median	10.95	10.92	9.77	7.18	6.99	6.51
Acc. (E_u) Mean	15.35	15.17	14.18	9.89	9.24	8.68
Speed coherence	1.46	1.17	1.08	1.99	1.30	0.97
Angle coherence	53.54	47.23	37.55	60.36	45.28	45.46

Table 1: Accuracies (in meters) [14] and further performance metrics of the deployed positioning methods, obtained at a hospital complex in Aarhus, Denmark.

4 ACTIVITY RECOGNITION APPLICATIONS AND METHODS CONSIDERED

HAR enables the realization of many context-aware and logistic applications, however, the nature of the environment and the complexity of human activities influence the recognition accuracy significantly. Understanding how variations in the input data affects different domains can improve the promising area of HAR. For the sake of concreteness, we focus in this paper on one of the prominent use-domains within applied HAR research, namely the smart hospital, which is a highly diverse and dynamic environment. A multitude of context-aware and logistic applications have been proposed for this domain, and we will focus on two such application and the respective underlying HAR formulations, namely Task Phase Recognition (TPR) and Indoor Transportation Mode Detection (ITMD). Both applications have been deployed and evaluated in the same environment, as have been the proposed indoor positioning algorithms—which provides the opportunity to evaluate the dependence between HAR performance and positioning methods in comparable scenarios for two different HAR applications. We will briefly describe proposed solutions to these two HAR problems, however, for a complete description we refer to the original papers [20, 27].

4.1 Task Phase Recognition

TPR has been proposed by Stisen et al. [27] to facilitate the (semi-)automatic coordination and communication between highly mobile workers in their daily routines. The methods proposed for TPR are evaluated on real-world sensor data collected through 4 days of manual observations of orderlies at a major university hospital. Central to the realization of such systems are the deployment, calibration and maintenance efforts [8], motivating the need to investigate how positioning accuracy influence the actual goal of TPR.

For illustration, we focus on TPR for orderlies, whose most common task is patient transport, which consists of the following four phases (1) *to task*, (2) *prepare patient* at the dispatching area, (3) *transportation* and (4) *handover patient* at the receiving area. The proposed TPR methods use machine learning classifiers commonly utilized in other activity recognition areas to recognize these phases. For features, Stisen et al. propose and investigate, among others, features extracted from positioning data—which is calculated using WC in

[27] and for which we herein also consider alternative positioning methods. The original best feature type combination includes direct line distances to the dispatching and receiving departments and historical proximity measures, which are both based on the underlying positioning system. This combination also proved superior compared to combinations with features using other input types, such as accelerometer data. In the original study, two different destination granularities were investigated, i.e. the accuracy of the dispatching and receiving locations. With department granularity, only the department boundaries are known and the middle of the department is used as the reference point. With room granularity, the exact room of the two destinations is known. Both granularities are also included in this study. To enhance the opportunities of generalizing TPR to other domains and use cases it becomes essential to investigate which positioning characteristics influence the recognition accuracy. In the following, we give a brief summation and categorization of said positioning-based features, and refer for a more detailed account to [27].

Live Proximity Tracking Features capture where a device or individual is relative to some contextual entity at the current moment in time. In our case these features are based on distances to the dispatching and receiving departments. We distinguish between four live proximity feature types —where each feature type includes as features standard statistical descriptions of the distances, as described in [27]. The four feature types then differ by first the distance used, either Direct Line (DL) or Shortest Path (SP) within the hospital's route network.

Historical Proximity Tracking Features capture whether a device or individual has been in the proximity of certain locations throughout the current history of a task. We adopt the proximity probability features described in [27] for both department and room granularity (referring to whether proximity to departments or to individual rooms–e.g. of start and end locations–is checked) however, we also introduce historical versions (H) of the live proximity tracking features. In the historical version, we simply remember the smallest average and minimum distances to the dispatching and receiving departments in the history of a task.

4.2 Indoor Transportation Mode Detection

Transportation mode detection (TMD) has received significant interest from the research community and has been applied in applications such as public transportation, environmental footprinting, and context-aware mobile assistants. Recently, TMD has been proposed for indoor scenarios [20], e.g., at hospitals or airports where a variety of vehicles is in common use. Prentow et al. [20] proposes methods which are also based on common machine learning classifiers for distinguishing between 6 different transportation modes: (1) *stationary* (2) *walking* (3) *scooter* (4) *bike* (5) *e-bus* (6) *e-bedpusher*. The article evaluates proposed features based on position estimates, however, it is concluded that these features do not perform very well, most likely due to the low accuracy of the WC positioning method. The best recognition accuracy in [20] is achieved by combing kinetic *time domain* features, i.e. features based on motion sensor data from worn or attached devices, with raw RSSI-based *WiFi* features, and does in other words not incorporate positioning based *spatial* features.

The immediate purpose of position tracking when utilized for ITMD is to estimate general motion characteristics such as foremost speed and to a lesser extent rate of turn, and acceleration patterns. Knowing these characteristics provides good separations for transport types such as being stationary from walking, and walking from transports by vehicle. Thus, the usefulness of a positioning system for ITMD is largely determined by how accurately one can derive such motion characteristics, foremost speed, from the provided positioning data. We investigate if utilizing better performing positioning methods yields better speed estimates and thereby better-performing spatial features. We evaluate the proposed positioning algorithms by calculating spatial features based on speed estimates between all consecutive position pairs in a window. Furthermore, an additional feature is also included in *S*, where the window is split into two temporal halves and the speed is then computed based on the average position in each half.

5 EVALUATION METHODOLOGY

The test-bed. To empirically evaluate impacts of positioning performance on HAR results, we used as main test bed a university hospital spanning 160,000m^2. At this hospital WiFi-based positioning and activity recognition are already used in daily operations as part of a software solution for hospital logistics. The solution was in large parts developed within a transdisciplinary research project, in cooperation with Systematic A/S. The hospital's already existing WiFi infrastructure includes 589 APs configured to send with constant emission power, spread evenly throughout the building complex, with 227 located on the ground floor, 206 on the 1st floor and 156 in the basement. WiFi measurements were collected from three different sources, i.e. from (devices kept or attached to) orderlies, patients and equipment.

Study limitations. Our evaluation profits from that several earlier studies were performed on data from the same real-world environment and real-world usage scenarios. A first study evaluated different WiFi positioning variants at the hospital, making use of extensive positioning ground-truth collections [14]. In terms of comparability and of reflecting real-world conditions of results across two chosen HAR tasks, TPR [27] and ITMD [20], two further studies were carried out, which also collected ground-truth on task phases and transportation modes, respectively. Focusing on real-world conditions in a hospital setting though comes with limitations in regards to the gathering of precise positioning ground truth—since such collection can interfere with and obstruct the natural work flow (and thus the real-world usage conditions) which we aimed to observe by following the orderlies in their daily work. Thus, for comparative positioning performance we rely on the first study [14], since when following orderlies we only gathered semantic destinations sporadically to confirm phase changes. Unless stated otherwise, results reported in the following are based solely on the orderlies' WiFi data collected by the network server; however, we will also report on the benefits of data fusions from all three sources and consider Wifi data collected by the clients as opposed to by the server.

Note, that a few tasks in the data set consist of incomplete data and that we will use the data with prior removal of incomplete tasks. We deem a task incomplete if no WiFi data is available for two or more complete phases in a task. All position estimates have been computed with a 10 second window size of WiFi measurements and 50 % window overlap. The same window segmentation is used for

TPR and we adopt the mixed window segmentation for ITMD presented in [20]. To ensure a consistent number of position estimate in each recognition window, we interpolate all estimate trajectories prior to the feature generation in the evaluations.

Evaluating recognition performance. To evaluate the performance of the features with different positioning methods we use the random forest classifier, since this classifier has been identified to produce the best results for the TPR as well as for the ITMD tasks [20, 27]. We also adopt the F_1-score as the primary evaluation metric for recognition accuracy, which is the harmonic mean of precision and recall. In line with previous HAR research [23] we report the weighted average of the F_1-scores of all classes:

$$Avg.\ F_1\text{-}score = \frac{\sum_{i=1}^{c} w_i * F_1^i\text{-}score}{\sum_{i=1}^{c} w_i}$$

where F_1^i-*score* is the F_1-*score* of the ith class and w_i is the number of samples of class i in the test dataset. To evaluate TPR we also adopt 10-fold cross validation and enforce that (data from) tasks are not split across folds and instead are allocated to only one fold each. This models real-world use-scenarios in a way that prevents unrealistically optimistic results, as often obtained which random folding of time-series data, see e.g. [26].

To assess whether some reported differences in the recognition distributions are significant, we employ McNemar χ^2-tests with Yates' correction [15]. More concretely, in our evaluation, this statistical test of paired nominal data combined with F_1-scores is used to determine whether an increase in recognition accuracy can be deemed significant. In line with empirical research in HAR and other domains, we will use p-values below 0.05 as indication of a significant difference in the distributions, i.e. rejection of the null hypothesis.

6 RESULTS

To illustrate the positioning performance differences for the positioning methods listed in Section 3, Fig. 1 depicts position estimate traces for a transportation task within the test bed hospital. An estimate's color refers to the task phase during which it was obtained: Yellow and red paddles represent position estimates during the phases *Prepare* (yellow paddles) and *Handover patient* (red), *To task* (green) and *Transportation* (yellow). Green and red pins mark the center of the *Dispatching* and *Receiving Departments*. Visibly, accuracy and coherence differ across the positioning methods—e.g., PF with RM (Fig. 1d) yields more accurate and coherent position estimates compared to, e.g., WC (Fig. 1a).

In the following, we present results showing how positioning performance influences the recognition accuracy for the TPR and ITMD problems. In section 6.1 we first present results showing how positioning performance influences various but single feature types for TPR, c.f. Section 4.1. In Section 6.2, we present respective results when combining different features types for TPR, and in Section 6.3 we analyze the effects of artificial degradation of positioning on the TPR recognition results. In Section 6.4, we present results for when incorporating positioning data for not only the orderlies but also patients and equipment. In Section 6.5, we present how positioning performance influence the recognition accuracy of the ITMD problem. Section 6.6 summarizes lessons learnt from the study.

			WC	TF(WC)	PF(WC)	RM	TF(RM)	PF(RM)
D	L	DL	72.51	73.43	74.03	74.66	75.48	77.06
		SP	68.90	70.68	71.71	72.46	72.82	74.13
	H	Prob.	58.42	58.93	59.09	59.94	57.59	57.01
		DL	58.81	55.76	54.31	57.27	57.59	52.74
		SP	52.01	53.46	53.81	57.17	59.01	53.12
R	L	DL	78.32	78.69	77.48	78.31	79.65	79.80
		SP	78.64	79.42	78.79	78.45	79.13	79.72
	H	Prob.	65.78	61.39	60.22	62.06	62.23	60.97
		DL	62.95	67.59	66.02	67.40	64.18	63.77
		SP	61.84	65.79	65.94	63.10	69.12	62.41

Table 2: F_1-scores (%) for TPR, when using either the live (L) or historical (H) proximity tracking feature type, and for both department (D) and room (R) granularity, respectively.

6.1 Assessing Impacts on Single Feature Types

In order to evaluate how the different positioning methods described in Section 3 impact individual feature types, Table 2 shows TPR recognition accuracies when training with single features types—from both the live proximity tracking and historical proximity tracking feature categories and for both department and room granularity.[1]

Rows 1 and 2 of Table 2 depict the result for the live proximity tracking features with department granularity. These results show a correlation with the positioning accuracies presented in Table 1, where, e.g., RM with PF yields both the best positioning as well as resulting recognition accuracy of 6.51m and an F_1-score of 77.06 respectively—while WC yields both the lowest positioning accuracy as well as recognition accuracy. Furthermore, filtering with PF always improves recognition accuracy significantly (according to the McNemar-Tests, c.f. Section 5), and filtered as well as non-filtered RM variants yield significantly higher recognition accuracies than the respective WC pendants. Such correlations between positioning and recognition accuracies are much weaker for historical proximity tracking features (with department granularity, see rows 5 through 9); e.g., RM with PF yields worse TPR recognition than RM alone. However, most often RM yields better TPR recognition than WC for historical proximity tracking features.

When considering not only department but room granularity for specifying destinations, as expected the recognition accuracy increases. Otherwise, the observations reported above for department granularity only partially hold when considering room granularity instead. Overall, the results suggest that the live proximity tracking features clearly benefit from increased positioning accuracy whereas the historical proximity tracking features do not to a similar extent. Furthermore, since using PF often worsens the results for the historical proximity tracking features, we suspect that for activity recognition tasks such smoothing, when applied to series of (sufficiently) accurate estimates can have a detrimental effect, eliminating the most accurate estimates.

6.2 Combining Feature Types for TPR

In this section, we analyze recognitions accuracies, as before but now for combining features of several types. The feature type combinations considered will consist of one contextual tracking and

[1] For the analysis of results in Table 2, we will focus on the F_1-score differences between positioning systems, and for and between department vs. room granularity, and refe the reader to a more detailed explanation of the features and their characteristics to [27].

(a) Weighted Centroid (WC) **(b) WC with Particle-Filtering** **(c) Radio Map (RM)** **(d) RM with Particle-Filtering**

Figure 1: Position estimates obtained during a single transportation task; phases of the task are differentiated by color.

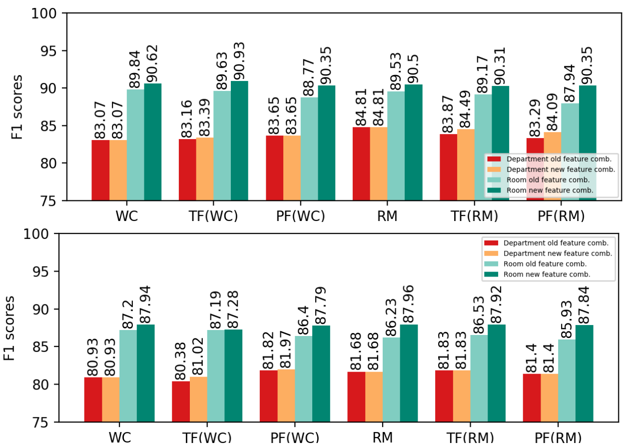

Figure 2: F_1-scores (%) for the original and new best feature type combinations for department and room granularity and for both 10s (top) and 4s (bottom) windows used for both positioning and recognition.

one historical proximity feature type, since the two categories utilize different aspects of the position estimates. Specifically, Figure 2 depicts F_1-scores for such combinations, and for 10s and 4s window sizes (used for both positioning and recognition), respectively. Furthermore, we list results for the best such feature combinations, according to using the original feature space described in [27] as well as the feature space as augmented for this paper.

Surprisingly, the increase in recognition accuracy observed when learning with live proximity tracking features alone does not necessarily propagate to combinations with historical proximity tracking features. This can partially be explained by the feature importances as determined by the random forest classifier we used: The single most important feature, when looking at the original best feature combination, is the dispatching department proximity sum (capturing the likelihood that the orderly has been at the dispatching department). Also, since the filtering methods did not improve the historical proximity features, but only the live tracking features, this effect may be too marginal to result in an overall improvement.

Choice of positioning methods for combining features. Another important observation from Figure 2 is that the optimal feature combination changes with the positioning method used: With increasing positioning accuracy, the benefits of the proximity probability features become smaller, i.e., comparatively it becomes more helpful to simply remember smallest distances to the destinations. Note, that it is not necessarily the same feature type combinations in the rows with the results for the new best combinations, whereas the original best combination is always the historical proximity probability feature combined with the direct line feature type.

The influence of the choice of features increases when the destination accuracy increases, see Figure 2 for results for room granularity and 10s windows. For all RM variants as well as PF(WC) there are significant gains in recognition accuracy by changing feature type combination. Furthermore, the observations made from the results in Table 2 and Figure 2 might suggest that different features have different 'best' positioning methods, i.e. positioning methods leading to best recognition results. One may conclude, that, it thus may be beneficial when combining features to combine also positioning methods: i.e., to use as input for each feature its respective best positioning method. However, when we ran all features combinations with respective best positioning methods we saw no gain in F_1-score compared to just using one positioning method.

Client-based positioning. Finally, the impact (for both HAR tasks, TPR and ITMD) of WiFi positioning architecture type is rather small: exchanging the WiFi data collected by the network for data collected from clients has seemingly negligible impact—the resulting alterations of the numbers in Figure 4 are less than one percentage point each and show no clear winner between the two architecture types.

6.3 Artificial Degradation

To assess the impact on HAR for a wider range of positioning performance levels than provided by the investigated positioning methods, c.f. Table 1, we also assessed the effects of (artificial) positioning degradation via addition of positioning noise. More specifically, we added white noise to position estimates by moving each of them by a random distance between 0, resp. $error_{min}$, and $error_{max}$ in a random direction, i.e. using a uniform distribution in the ring around the estimate defined by those two radii.

To illustrate the effects of the degradation on the positioning, the resulting positioning performance measures, as listed in Table 1, for the six investigated positioning variants, are given in Table 3 for the weighted centroid (WC) positioning variant degraded by noise of different level and coherence. Since our results suggest that positioning coherence is often equally as important for HAR applications as positioning accuracy, we also simulate different levels of coherence by applying a moving average of varying length k to the noise vectors, comprised of direction and angle. The noise vector for each new estimate is obtained as the average over one newly created vector and the $k-1$ vectors created for the previous estimates. Fully incoherent noise vectors we obtain for $k=1$. With $k=n$ we denote the extreme case, that the same noise vector is applied to all estimates within a

Performance measure	10m	40m	80m	10-80m
Acc. (E_u) Mean	16.87 / 17.33 /17.49 / 17.52	25.10 / 27.31 / 27.85 / 27.97	39.32 / 44.04 / 45.66 / 46.07	42.39 / 48.40 / 50.39 / 51.05
Speed Coherence	2.71 / 1.76 / 1.48 / 1.46	10.32 / 4.84 / 1.98 / 1.46	21.23 / 9.70 / 3.11 / 1.46	22.82 / 10.70 / 3.36 / 1.46
Angle Coherence	71.75 / 63.49 / 56.30 / 53.54	83.33 / 76.51 / 64.69 / 53.54	86.06 / 81.61 / 70.89 / 53.54	85.94 / 81.56 / 71.69 / 53.54

Table 3: Accuracy, speed and angle coherence for WC degraded by adding varying level of noise and coherence (k=1,4,16,n).

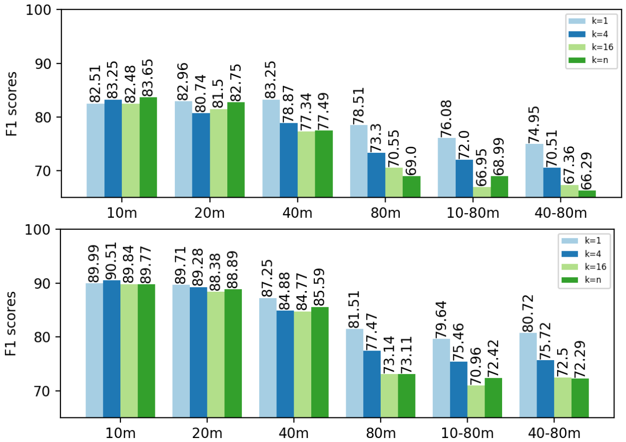

Figure 3: TPR results as F_1-scores (%), for varying levels of positioning degradation and smoothing; for department granularity (top) and room granularity (bottom), using the original best feature type combination.

recorded trajectory. Table 3 illustrates the growth of inaccuracies as well as of speed and angle incoherences with growing noise level; it also illustrates the how speed and angle coherence improves with growing k, i.e. with growing coherence of the noise applied.

Figure 3 presents F_1-scores (%) for the TPR scenario when WC positioning with varying degradation is used as input. As for non-degraded positioning, c.f. Table 1, the results provide further indication that diminished positioning accuracy has surprisingly small impacts on TPR accuracy. This backs our intuition that in TPR for transport tasks the most valuable observation data aspect is the overall travel pattern, and that positioning accuracy is sufficient as long as it can detect area-leave and -enter events somewhat accurately and with limited latency.

Coherence of the noise, and thus of the degraded positioning, has no or even a positive impact on TPR accuracy for minor noise levels, i.e. for $error_{max} \leq 20m$, while for larger noise levels, a strong negative impact can be observed. Again, the results suggest that to (learn to) differentiate task phases correctly hangs on whether some positions of higher accuracy, even when sporadic, i.e. occurring at low frequency, exist. When smoothing the (higher-level) noise via moving averages, these frequency of higher-level accuracy positions may drop below a threshold for which (especially: area-leave or -enter) events are incorrectly detected.

6.4 Fusing Orderly, Patient and Equipment Data

The results reported in Section 6.2 suggest that the TPR learning, resp. recognition, profits from additional information such as combinations of several feature types, and switching from lower department to higher room granularity. This additional information then can also make up for less accurate positioning input. In this section, we investigate the benefits of using input position data not only from orderlies also from patients and from equipment. Respective results

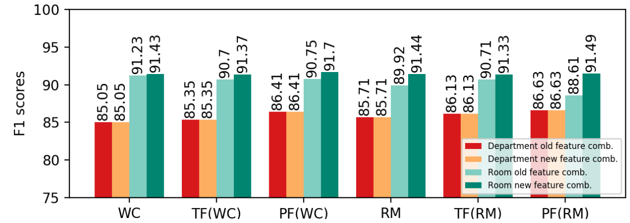

Figure 4: F_1-scores (%) for original and the new best feature type combinations for department and room granularity when using three positioning data sources.

are given in Figure 4 for the original and the new best feature type combinations for TPR.

Generally, the recognition F_1-score is boosted by a few percentage points by the additional input data. When comparing with the results for only orderly positioning input, see Figure 2, the filtering methods now also improve the recognition accuracy when combining features for department granularity. This is likely because the improvements on the live proximity tracking features from PF are important enough to also improve the combined recognition accuracy. Even more, the tracking of additional targets, i.e. patients and equipment, provides similar or better gains (1.98pp by adding multiple positioning sources with WC, i.e. the results in Figure 2 and 4, respectively) than switching positioning system from the one with worst to the one with the best F_1-score (1.74pp in Figure 2 alone).

When comparing position estimates from several sources, coherence likely becomes more important, such that co-moving of both the orderly, patient and equipment can be inferred. The single most important feature for the patient is also not the dispatching department probability sum, as is the case for the orderly, i.e. it is likely more important to know if the patient is close to the orderly than whether the patient is at the dispatching department. Furthermore, as stated in [14] the RM method is less coherent, which can explain the smaller absolute differences in F_1-score between the WC and RM methods, when compared to the results in Figure 2. The choice of combination of features types can often be adjusted to maintain a stable recognition accuracy—since, as previously, the best feature type combination depends on the positioning method used.

6.5 Indoor Transportation Mode Detection

The original investigation into ITMD by Prentow et al. [20] reported F_1-scores of 79.73% and 85.96% for using accelerometer based time domain features alone, and for combined with raw WiFi signal strength features, respectively. The latter feature combination was evaluated as superior by Prentow et al. and serves as a benchmark for our evaluation. Table 4 depicts the ITMD results for spatial positioning-based features alone and in various combinations with time domain features and WiFi signal strength features for all the proposed positioning methods. We also included results where we lowered the window size for positioning from 10s to 4s.

While the spatial features when utilized alone yield low F_1-scores for all investigated positioning methods, the spatial features still

		WC	TF(WC)	PF(WC)	RM	TF(RM)	PF(RM)
10s	S	69.15	68.85	68.86	66.06	69.19	68.13
	T, S	83.93	83.90	84.54	83.04	83.25	84.09
	T, W, S	86.37	85.87	86.31	86.50	86.34	86.74
4s	T, W, S	86.63	87.20	87.16	86.96	86.65	87.00

Table 4: F_1-scores (%) for ITMD of spatial features alone and in various combinations with time-domain accelerometer (T) and WiFi (W) features for all the proposed positioning methods and both 10s and 4s window sizes for positioning only.

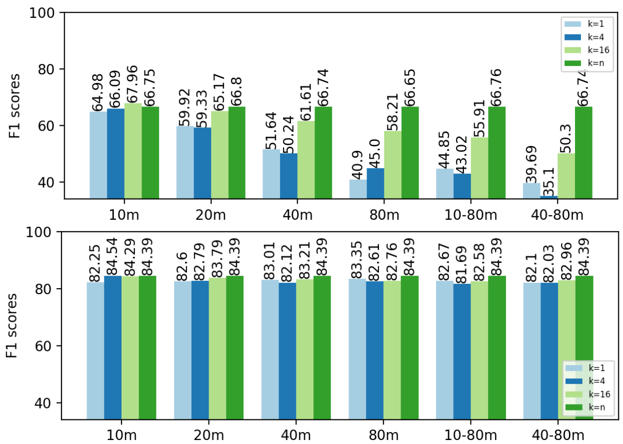

Figure 5: F_1-scores (%) for ITMD using spatial features alone (top) and together with time-domain accelerometer features (bottom), for WC positioning and for added noise of varying level and coherence.

prove useful when combined with other feature types: e.g. when using TF(WC) for positioning, the spatial features, c.f. Section 4.2, improves the benchmark result by 1.26*pp* to 87.20% when added to the original best feature type combination. However, the results of Table 4 also show that the spatial features cannot without losses replace the features based on raw WiFi signal strengths, which are designed to be a proxy for speed. Furthermore, the gain in positioning performance has little to no impact on the recognition accuracy. The results suggest that incoherence of a positioning method is more harmful to ITMD performance than positioning inaccuracy, because incoherence will have a more direct impact on the speed estimation error. The explanation aligns with what can be observed when artificially increasing and varying positoning inaccuracy and incoherence, by adding noise (as detailed in Section 6.3); see Figure 6.5. For incoherent noise ($k = 1$), the TPR performance drops with increasing noise level; but the more coherent the added noise is (i.e., the larger k) is, the less affected is TPR performance by that noise (even when of high level).

However, Figure 6.5 (bottom) also reveals that when combining positioning-based features with suitable other (here: accelerometer) features, learners can also make use of poor positioning input: For all levels and coherences of added noise, the addition of spatial features to the accelerometer-based features results in very similar slight increases (i.e. around 2.5pp-4.5pp) in F_1-score.

Somewhat surprisingly, when fusing all three feature types, lowering the window size for positioning can increase the recognition accuracy slightly for ITMD. While the overall positioning accuracy is likely better for large window sizes, the averaging of signal strengths might hide speed (change) patterns that provide useful information for ITMD.

6.6 Lessons and Guidelines for ITMD and TPR

Overall, our results reveal that the impact of positioning performance on HAR applications such as TPR and ITMD is not easy to capture. When only considering positioning accuracy as the performance measure: the correlation between positioning performance and recognition accuracy changes based on the setting of the recognition problem, and generally the impacts of positioning accuracy tend to be surprisingly small. One of several explanations to this, is that higher positioning accuracy does not necessarily imply higher speed coherence and angle coherence, as shown in Table 1.

For TPR, while improving positioning accuracy occasionally has a positive effect on recognition accuracy, e.g. when using the live proximity tracking features, there are several scenarios in which this improvement is mitigated by other factors, such is additional targets being tracked, feature type combinations and discrepancies between the different positioning performance measures. Only when adding significantly larger accuracy errors via artificial degradation a clear impact of positioning accuracy could be observed.

For ITMD, we saw no notable recognition improvement from increased positioning accuracy. The artificial degradation have shown that the lack of improvement in speed and angle coherence is one of the reasons no clear improvements could be observed. Furthermore, since acceleration patterns are important to distinguish the different transportation modes, position estimates are not the only source of information in the optimal ITMD solution. The dependency on an additional data type also mitigates the effect of better positioning performance.

Overall, the impact of measures for the coherence of positioning, such as of speed and angle, c.f. Section 3, seem more pronounced than the impact of positioning accuracy. Interestingly, manufacturing coherence via smoothing proved beneficial only for ITMD, while for TPR simple filtering may even harm the recognition accuracy. As the effect of positioning performance, and the best choice for a positioning solution depends on many factors and inter-dependencies and seems to be application-dependant as well, we derived below a set of concrete guidelines regarding positioning and HAR system choices for the tasks of TPR and ITMD before we discuss some generalized lessons in the next section.

- For TPR, feature design is generally more important than positioning accuracy.
- For TPR, effort is better spent gaining room level information instead of implementing better positioning.
- For TPR, tracking additional targets (such as patients and equipment) and utilizing that data leads to higher recognition accuracies, owing especially better live proximity tracking feature performance for department granularity.
- For TPR, smoothing the positioning may be harmful–likely since it may prevent accurate and timely hints about leaving or approaching of rooms. Relatedly, for higher destination granularity (e.g. room), historical proximity tracking features become the most valuable feature type.
- For ITMD, enlarging positioning window sizes may be harmful, this may average-out and thus hide important speed patterns.

- For ITMD, smoothing and high coherence of positioning is benefitial. If only low coherence is achievable, more reliable speed indicator than the positioning data may be better suited to boost recognition accuracy, e.g. raw WiFi-RSSI changes.
- If the choice of positioning method is restricted, features can be designed which reflect the uncertainty of the positioning and thereby improve recognition accuracy. An example of this are the proximity probability features presented in [27].

7 DISCUSSION

While TPR and ITMD seemingly are rather specific HAR applications, we believe the presented results generalize to other use domains and HAR tasks with similar usage of position estimates. In TPR, position estimates are used through repeated comparisons with known reference points in the environment and the position estimates is the primary source of information. In ITMD, position estimates are used through comparisons with other position estimates and accelerometer data is an equally important source of information. These considerations are important when assessing whether the results generalize to another HAR task.

The system for improving logistics at hospitals, born in part out of the research reported on here, attracts great demand and is employed successfully in many hospitals, and is used for a variety of sensing- and HAR-based services, including tracking and guiding of personnel, visitors, assets, and vehicles. Interestingly, the sensing-based functionalities needed for all these services are mostly variants of the tasks investigated here, task phase and transport mode recognition, as well as basic positioning and proximity services. Furthermore, generalization also seems possible to other larger indoor and mixed domains where logistic operations are complex. Such domains include airports, train stations, large warehouses, factories and farms.

In the following, we will discuss general themes of this investigation. We invite researchers to debate and extend them, and to add to the investigations herein in regards to, e.g., further HAR and positioning-based applications, solutions, evaluation procedures and environments.

More Data or Better Learning vs. More Accurate Positioning. It is an often cited principle in the machine learning community that "more data beats a clever algorithm" [3]. The results presented in this paper show that the gains in recognition accuracy achieved by improving positioning performance are relatively small. This may suggest that the effort in developing and deploying improved positioning methods is often better spend on gathering more, and more diverse, empirical data: Benefitial is not only a larger training data base, as expected, but specifically also the gathering of data from additional sources, e.g. in the hospital domain through the tracking of not only orderlies but also patients equipment proved useful (as shown for TPR), or the equipping of such entities with additional sensing modalities such as motion sensors (as shown for ITMD).

Seemingly, also efforts spent on feature design, choice, and adaption seems to yield higher (and cheaper) returns than the pursuit of higher positioning accuracy. Prominently, the reflection on and categorization of positioning-based features may aid better HAR design also for HAR problems beyond the ones investigated here. We believe this to hold, e.g., for the distinction between live and historical proximity tracking feature types and for the reflections on proxies for speed, ranging from raw, unfiltered signal strength data to smoothed positioning traces.

For future research, we believe that similar studies on larger data sets will be fruitful—especially investigating whether and to which extent increases in data amount and variety translates to increases in recognition accuracy, and to compare these increases with those obtained through increased positioning performance. Generally and in line with current real-world requirements as motivated in [14], which states that development and maintenance efforts ideally should be minimized, it would be interesting to develop an understanding of how effort is best spent, since this could assist in the design process of future real-world positioning and HAR systems as well as deployment guidelines and procedures.

Positoning Performance Indicators Beyond Accuracy. Our results indicate that choosing a positioning method with a higher positioning accuracy does not necessarily result in a similar increase in activity recognition accuracy. However, impacts of positioning accuracy, as well as of other positioning performance characteristics, on recognition accuracy have been observed.

The cause-effect relations between positioning accuracy and activity recognition performance is often mitigated by additional factors. For instance, while WC is evaluated to have the lowest accuracy of the proposed positioning methods, it still provides competitive results, e.g., through the design of probability proximity based features, which can account for the low positioning accuracy. Furthermore, it can also be argued that WC describes the variations in the underlying measurements more directly, which then benefits the classifier, since large variations might happen more often during transport.

More importantly, though, our results also suggest that sometimes i) other performance measures, such as positioning coherence, have a higher impact on HAR accuracy than positioning accuracy does, and ii) for some HAR applications, boosting positioning coherence via smoothing techniques can boost HAR accuracy. The latter though seems dependent on the HAR application at hand, and to hold primarily for HAR scenarios where positioning is used to extract features capturing motion characteristics such as speed (as in ITMD, where transport modes can be distinguished by speed) rather than capturing absolute positioning (as in TPR, where proximity to certain areas is most relevant to detect).

For such reasons, comparing our results with future work on real-world evaluations, using additional positioning methods and use scenarios, seems worthwhile. We propose that future research should focus more on evaluating direct causation, since the results of this evaluation have shown that the relation between the common evaluation metrics in two separate task domains can be unclear and unexpected.

Interdependences between features and positioning methods. The evaluations of this paper, especially the results for TPR, provides evidence that there are not best features and best positioning systems, but rather suitable and less suitable combinations of the two: For instance, smoothing of positioning traces leads to higher accuracy when using live proximity tracking feature types; whereas smoothing leads to no or negative gains when using historical proximity tracking feature types instead. This also implies that when, e.g., changing the positioning system used, the effect of HAR performance is hard to predict, and may also profit from a re-adjustment of the feature space and the learning system. As a resulting rule of thumb, if the increase in positioning performance is not sufficiently large, no or only little increase in HAR recognition is to be expected, unless the HAR system itself is re-investigated and -adapted.

The results presented in this paper also show that the individual gains that feature types receive from increased positioning performance is not additive: The significant recognition gains that can be achieved for individual feature types by improving positioning does not necessarily propagate when combining feature types. For instance, RM provides significantly lower F_1-scores than the smoothed positioning methods when learning with live proximity tracking features alone. However, for department granularity, RM proves superior when combining the two feature categories; but when adding additional data sources (in the form of additional tracked targets) the picture changes once again. In other words, how well features and positioning methods complement each other is not evident from learning with single feature types.

8 CONCLUSION

In this paper we have presented an evaluation of the impacts of indoor positioning performance on activity recognition, specifically on two HAR tasks, namely task phase recognition and indoor transportation mode detection. The results reveal that improved positioning provides increased recognition accuracy, though to a lesser extent than anticipated and only with diminishing returns. The results also indicate that positioning performance measures other than the traditionally reported discrete accuracy, specifically measures for positioning coherence, may provide better indicators for resulting HAR performance. Generally, the results suggest that it might not always be beneficial to opt for a more accurate positioning method, when taking effort into account, since less accurate or precise positioning can support the intended use. Instead, gains in HAR may be achieved more easily by better utilization of existing positioning solutions and by careful adaption of the HAR learning system. We hypothesize that similar investigations of the direct relation between indoor positioning performance and the actual utilization benefits for positioning-dependent applications will assist in identifying future challenges as well as best practices and hence ease future deployments.

REFERENCES

[1] Paramvir Bahl and Venkata N Padmanabhan. 2000. RADAR: An in-building RF-based user location and tracking system. In *Proc. IEEE INFOCOM*, Vol. 2.

[2] Jakob E Bardram, Afsaneh Doryab, Rune M Jensen, Poul M Lange, Kristian LG Nielsen, and Søren T Petersen. 2011. Phase recognition during surgical procedures using embedded and body-worn sensors. In *Proc. IEE PerCom 2011*. IEEE.

[3] Pedro Domingos. 2012. A few useful things to know about machine learning. *Commun. ACM* 55, 10 (2012), 78–87.

[4] Frédéric Evennou and François Marx. 2006. Advanced integration of WiFi and inertial navigation systems for indoor mobile positioning. *Eurasip journal on applied signal processing* 2006 (2006), 164–164.

[5] Ramsey Faragher and Robert Harle. 2015. Location fingerprinting with bluetooth low energy beacons. *IEEE Journal on Selected Areas in Communications* 33, 11 (2015), 2418–2428.

[6] Jesus Favela, Monica Tentori, Luis A Castro, Victor M Gonzalez, Elisa B Moran, and Ana I Martínez-García. 2007. Activity recognition for context-aware hospital applications: issues and opportunities for the deployment of pervasive networks. *Mobile Networks and Applications* 12, 2-3 (2007), 155–171.

[7] Brian Ferris, Dirk Haehnel, and Dieter Fox. 2006. Gaussian processes for signal strength-based location estimation. In *In proc. of robotics science and systems*. Citeseer.

[8] Thomas Riisgaard Hansen, Jakob E Bardram, and Mads Soegaard. 2006. Moving out of the lab: deploying pervasive technologies in a hospital. *IEEE Pervasive Computing* 3 (2006), 24–31.

[9] Thomas King, Thomas Haenselmann, and Wolfgang Effelsberg. 2007. Deployment, calibration, and measurement factors for position errors in 802.11-based indoor positioning systems. In *Location-and Context-Awareness*. Springer, 17–34.

[10] Mikkel Baun Kjærgaard. 2007. A taxonomy for radio location fingerprinting. In *Location-and Context-Awareness*. Springer, 139–156.

[11] Mikkel Baun Kjærgaard, Mads Vering Krarup, Allan Stisen, Thor Siiger Prentow, Henrik Blunck, Kaj Grønbæk, and Christian S Jensen. 2013. Indoor Positioning using Wi-Fi–How Well Is the Problem Understood?. In *Proc. IPIN*.

[12] Heli Koskimäki, Ville Huikari, Pekka Siirtola, Perttu Laurinen, and Juha Röning. 2009. Activity recognition using a wrist-worn inertial measurement unit: A case study for industrial assembly lines. In *Control and Automation, 2009. MED'09. 17th Mediterranean Conference on*. IEEE.

[13] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. 2011. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter* 12, 2 (2011), 74–82.

[14] Andreas Mathisen and Søren Krogh Sørensen. 2016. A Comparative Analysis of Indoor WiFi Positioning Algorithms at a Large University Hospital. In *Proc. IPIN*. IEEE. Preprint: http://tinyurl.com/zgct2w6.

[15] Quinn McNemar. 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12, 2 (1947), 153–157.

[16] Lionel M Ni, Yunhao Liu, Yiu Cho Lau, and Abhishek P Patil. 2004. LANDMARC: indoor location sensing using active RFID. *Wireless networks* 10, 6 (2004), 701–710.

[17] Nicolas Padoy, Diana Mateus, Daniel Weinland, M-O Berger, and Nassir Navab. 2009. Workflow monitoring based on 3d motion features. In *Computer Vision Workshops (ICCV Workshops)*. IEEE.

[18] Siddika Parlak, Ivan Marsic, and Randall S Burd. 2011. Activity recognition for emergency care using RFID. In *Proc. 6th Intl. Conf. Body Area Networks*. ICST.

[19] Francesco Potortì, Paolo Barsocchi, Michele Girolami, Joaquín Torres-Sospedra, and Raúl Montoliu. 2015. Evaluating indoor localization solutions in large environments through competitive benchmarking: The EvAAL-ETRI competition. In *Proc. IPIN*. IEEE.

[20] Thor Siiger Prentow, Henrik Blunck, Mikkel Baun Kjærgaard, and Allan Stisen. 2015. Towards indoor transportation mode detection using mobile sensing. In *Proc MobiCase*. Springer.

[21] Thor S Prentow, Antonio J Ruiz-Ruiz, Henrik Blunck, Allan Stisen, and Mikkel B Kjærgaard. 2015. Spatio-temporal facility utilization analysis from exhaustive WiFi monitoring. *Pervasive and Mobile Computing* 16 (2015), 305–316.

[22] Nissanka B Priyantha, Anit Chakraborty, and Hari Balakrishnan. 2000. The cricket location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking*. ACM, 32–43.

[23] Hesam Sagha, Sundara Tejaswi Digumarti, José del R Millán, Ricardo Chavarriaga, Alberto Calatroni, Daniel Roggen, and Gerhard Tröster. 2011. Benchmarking classification techniques using the Opportunity human activity dataset. In *Systems, Man, and Cybernetics (SMC)*. IEEE.

[24] Dominick Sanchez, Monica Tentori, and Jesús Favela. 2008. Activity recognition for the smart hospital. *Intelligent Systems, IEEE* 23, 2 (2008), 50–57.

[25] Thomas Stiefmeier, Daniel Roggen, Georg Ogris, Paul Lukowicz, and Gerhard Tröster. 2008. Wearable activity tracking in car manufacturing. *IEEE Pervasive Computing* 2 (2008), 42–50.

[26] Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind Dey, Tobias Sonne, and Mads Møller Jensen. 2015. Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition. In *Proceedings of the 13th ACM Conference on Embedded Network Sensor Systems (ACM SenSys)*. ACM, 127 – 140.

[27] Allan Stisen, Andreas Mathisen, Henrik Blunck, Mikkel Baun Kjærgaard, and Thor Siiger Prentow. 2016. Task phase recognition for highly mobile workers in large building complexes. In *Proc. PerCom*. IEEE.

[28] Roy Want, Andy Hopper, Veronica Falcao, and Jonathan Gibbons. 1992. The active badge location system. *ACM Transactions on Information Systems (TOIS)* 10, 1 (1992), 91–102.

[29] Jamie A Ward, Paul Lukowicz, Gerhard Troster, and Thad E Starner. 2006. Activity recognition of assembly tasks using body-worn microphones and accelerometers. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 28, 10 (2006), 1553–1567.

[30] Lauri Wirola, Tommi Laine, Jari Syrjärinne, et al. 2010. Mass-market requirements for indoor positioning and indoor navigation. In *Proc. IPIN*. IEEE.