A Review and Taxonomy of Activity Recognition on Mobile Phones

Ozlem Durmaz Incel · Mustafa Kose · Cem Ersoy

Abstract The release of smart phones equipped with a rich set of sensors has enabled human activity recognition on mobile platforms. Monitoring the daily activities and their levels helps in recognizing the health and wellness of the users as a practical application. Mobile phone's ubiquity, unobtrusiveness, ease of use, communication channels, and playfulness make mobile phones a suitable platform also for inducing behavior change for a healthier and more active lifestyle. In this paper, we provide a review on the activity recognition systems that use integrated sensors in the mobile phone with a special focus on the systems that target personal health and well-being applications. Initially, we provide background information about the activity recognition process, such as the sensors used, activities targeted, and the steps of activity recognition using machine learning algorithms, before listing the challenges of activity recognition on mobile phones. Next, we focus on the classification of existing work on the topic together with a detailed taxonomy. Finally, we investigate the directions for future research.

O. D. Incel (✉) · M. Kose · C. Ersoy
Department of Computer Engineering,
Bogazici University,
Bebek 34342, Istanbul, Turkey
e-mail: ozlem.durmaz@boun.edu.tr

M. Kose
e-mail: mustafa.kose@boun.edu.tr

C. Ersoy
e-mail: ersoy@boun.edu.tr

Keywords Mobile phone sensing · Human activity recognition · Classification

1 Introduction

Research on human activity recognition that involves the use of different sensing technologies provides a great potential for personal health systems by monitoring the daily activities and, hence, the wellness and health status of their users. Especially, the release of smart phones equipped with a rich set of sensors together with their ubiquity is a key enabler for the adoption of personal activity recognition systems on mobile platforms by the masses.

To briefly go over the history of such systems, vision sensing using cameras has been the focus of early research in the activity recognition domain [77]. More recently, inertial sensing, using movement-based sensors that can be attached to the user's body has been investigated [15]. In fact, today, mobile phones can act as an activity recognition platform. In the early studies, GSM signals were used to infer the basic transportation modes of the users [92], for instance, to understand whether a user is in a vehicle or is in stationary state. Also, in the early versions, activity recognition systems that use on-body wearable sensors as peripherals that can be connected to a mobile phone were utilized [21]. Currently, smart phones that are integrated with a rich set of sensors, such as accelerometer, gyroscope, GPS, microphone, camera, proximity and light sensors, Wi-Fi, and Bluetooth interfaces, provide a suitable platform for personal activity recognition systems.

Not only the inclusion of new sensors and their ubiquity but also their unobtrusiveness, zero installation cost, and ease of usability make activity recognition on mobile

phones more attractive. Compared to using on-body sensors, mobile phones do not disturb or limit the activities of the users, do not require to carry or attach external devices on the body, and do not require the installation or calibration of the sensors. Mobile phone-based activity recognition also eliminates the installation of cameras and the limitation of working indoors compared to vision-based sensing systems.[1] Besides these facts, a survey showed that a mobile phone is considered as essential when leaving home besides a key and a wallet, and this reveals that mobile phones naturally fit people's lives without disturbing them [11].

In fact, despite these advantages, activity recognition on mobile phones also faces challenges such as the battery limitation of the phones, limitations in processing and storage compared to more powerful stations, and human behavior, such as different use of phones by different people, that need to be investigated for the realization and further adoption of these systems.

From the perspective of personal health and well-being systems, the fundamental contribution of activity recognition on smart phones can be monitoring the physical health and well-being of the users by screening the physical activities, such as the transportation modes, locomotion, and sports activities, through inertial sensors, mainly through accelerometers. For instance, by monitoring the activity levels, daily energy expenditure and physical wellness can easily be calculated. Alternatively, the mobile phone can be used as a fitness coach, for instance, for the rehabilitation of patients with diseases, such as Parkinson's disease. The patient can be given a program with a set of activities, their duration, and sequence, and the phone can detect what the user is currently doing and/or if he is performing the activities correctly with a correct sequence. Moreover, by following the daily activities, the routines and behavior of the users can easily be learnt by the activity recognition systems, and, in case of a drift from these routines, the users can be warned about these changes. Especially considering the elderly, these drifts may be indicators for certain diseases, such as Alzheimer's disease or dementia, and informing caregivers may be crucial. Together with accelerometers, GPS can also provide very useful information for monitoring the activities including also the location information. For instance, again considering the elderly, following their daily trajectories and identifying drifts in these trajectories can reveal important findings about their mood and

cognitive well-being [42]. Other sensors such as the microphone, providing ambient sound information, camera, Bluetooth interface, providing information on social interactions, Wi-Fi and cellular radios, providing location information, and proximity and light sensors, providing ambient information, can enrich the activity recognition for personal health and well-being by providing additional context information.

Besides using the sensing functionality, smart phones are also convenient interaction platforms for persuasive applications to motivate healthier behavior because of their ubiquitous presence, communication channels, and playfulness together with their perceived role as a personal and trusted technology [6, 75] as investigated in [33, 35].Virtual companions [12], games [78], and social networks [99] can be used as the means to interact with the users in mobile applications in order to persuade them for changing their sedentary or unhealthy habits and even lifestyle for healthier behavior [27].

In this paper, we review the activity recognition systems that use integrated sensors in the mobile phone with a particular focus on the systems that target personal health and well-being applications. Our aim is to provide an extensive survey on the topic and to bring the researchers not working in the field quickly up to date about the state of the art, opportunities, challenges, and future topics on activity recognition using mobile phones. To the best of our knowledge, although surveys on activity recognition using wearable sensors exist [13, 54, 79], a survey on mobile phone sensing [51] and on the classification algorithms for activity recognition on smart phones [5], there is no extensive survey on activity recognition on mobile phones including a taxonomy of existing work especially focusing on the issues of health and well-being.

Our contributions are to present a detailed survey on the topic, provide a taxonomy of existing work, and investigate the open issues in this domain. We start with a background information on what types of sensors used, what kind of activities are targeted, application domains of activity recognition on mobile phones, and how activity recognition systems are or can be used in the field of personal health and well-being. We focus on the activity recognition process and explain the details of used techniques. We particularly explain the steps of the activity recognition using supervised machine learning algorithms. Then, we elaborate on the performance evaluation of the used techniques explaining the performance metrics utilized and experimentation of proposals. The next topic we focus on is the challenges of activity recognition on mobile phones and the possible solutions proposed in the literature to overcome these challenges. In the second part, we investigate the proposed solutions on activity recognition on mobile phones. We classify the solutions according to the sensors used and off-line versus online classification methods together with

---

[1] We admit that camera-based systems are more powerful in detecting a more detailed context for the user, and phone-based activity recognition may not replace the camera-based solutions in settings where each action is required to be detected. Here, our consideration is about the set of the activities that can be detected with the set of sensors available on the phones, and the camera still can be used for detailed context recognition.

presenting a taxonomy of existing work from different aspects such as the activities targeted, type of classification techniques used, and performance testing. In the last part, we provide a list of open issues and directions for future research.

The remainder of the paper is organized as follows: in Section 2, we provide the background information on activity recognition using mobile phone sensors, such as the types of sensors used, activities detected, and application domains. In Section 3, we focus on the process of activity recognition particularly on the steps of machine learning techniques and performance metrics. Section 4 is about the research challenges in the field of activity recognition on mobile phones, whereas Section 5 includes the taxonomy of existing work. In Section 6, we discuss the open issues and comment on possible future research directions. Finally in Section 7, conclusions are drawn.

2 Background

In this section, first we provide the background information on the sensors available on mobile phones that are or can be used for activity recognition purposes and next focus on the activities that can be inferred using these sensors.

2.1 Sensors

Although today's mobile phones are powerful devices with their computational capabilities and richer functionality with the integrated sensors, they still act as communication devices from the user's point of view [51]. The next step should be to utilize the potential of using the mobile phone as active assistant devices in supporting users' daily activities, and the main enablers to realize this step are the integrated sensors on the mobile phone and the improving computational capabilities of today's mobile phones.

Figure 1 shows an example set of sensors available on current smart phones: the conventional sensors such as the cellular radio, Wi-Fi radio, Bluetooth radio, microphone, cameras, and GPS, and newer sensors such as the accelerometer, gyroscope, compass, light and proximity sensors. It is expected that in the near future there will be more and more sensors integrated on the mobile phone to support a diverse set of applications. For instance, humidity and gas sensors can easily be integrated to infer more information about the user's context. Sensors providing health information, such as blood pressure or heart beat rate, may not be useful when embedded on the phone itself since they require contact with the user's skin, but they are already used as peripherals that can be connected to a mobile phone. In fact, there are applications that can measure the blood oxygen saturation and heart rate using the

Fig. 1 Example sensors available on smart phones

cameras available on the smart phones [1]. It is clear that, as the technology, particularly in microelectromechanical systems, improves and the research on mobile phone sensing matures, there will be new applications requiring more and more sensors either embedded on the phones or used as peripherals.

As a fundamental sensor on a mobile phone, the *radio for cellular communication* enabled the first ubiquitous applications for coarse-grained context recognition. By using the connection information between the radio and the cell-tower, it is possible to locate the user, such as when the user is at home. Although this does not give a fine-grained activity information, it provides location-based information to give a clue about what the user may be doing. Besides such a location-based service, signal fluctuation information between the radio and the cell tower has been used to predict the user's mode of transportation, such as walking, driving a vehicle, or in stationary state [92].

Besides the radio for cellular communication, *Bluetooth* and *Wi-Fi radios* can also be used as sensors for context and activity recognition. For instance, in the "Reality Mining Project" [30], interactions between or co-location of Bluetooth radios on the mobile phone were used to infer social interactions between phone users. Moreover, fluctuations in Wi-Fi signals can be used to locate the user, such as in the classroom or attending a meeting, and again to predict the user's mode of transportation [71].

The *microphone* can also be used for activity recognition purposes by collecting audio during a user's daily activities, such as being in a conversation and being in a noisy environment [65]. Similarly, the *camera* can also provide a rich set of context information. One example application, EyePhone [69], uses the camera to detect activities such as tracking the user's eye movements to start the applications on the phone. *GPS* is another powerful sensor for tracking the location of the mobile phone and the speed of movements similar to the cellular and Wi-Fi radios but with a more fine-grained information of location.

The *three-axis accelerometer* is one of the most effective sensors for activity recognition purposes on mobile phones. Although it was integrated in the mobile phone with the objective of enhancing the user experience by changing the orientation of the display according to the orientation of the phone held by the user [51], it can be used for activity recognition by inferring the user's movements, such as walking, standing, running, sitting, and even falling [98]. In fact, activity recognition using inertial sensors has been an active field of research [13], and recently, using accelerometers on the mobile phone is receiving a lot of attention from the research community. In Section 5, we summarize the example studies using accelerometers for activity recognition. Similarly, the *gyroscope* and *compass* can be used for activity recognition by measuring the orientation of the mobile phone and enhancing the results inferred with an accelerometer.

Proximity and *light* sensors are also the set of sensors embedded in phones for enhancing the user experience. For instance, when the proximity sensor detects that the user holds the phone close to her face, the keys are disabled or similarly the light sensor adjusts the brightness of the screen. For activity recognition purposes, they can be used together with other sensors to infer more accurate activity information. For instance, the light sensor may provide information about the environment of the user, such as being in a dark environment, and give prior information about the set of the activities that can be carried in a dark environment, such as sleeping but not reading a book.

It is also possible to use a combination of sensors to infer more detailed information about the user's activities. We will elaborate on such examples in Section 5.4, but to mention, for instance in [85], both the GPS and accelerometer information is collected to infer user's movements. GPS data can be used to identify if a user is walking or in a vehicle, but it is difficult to identify whether a user is running or biking by just looking at the speed. Combined with the accelerometer data, it can provide more fine-grained activity information.

2.2 Activities

In the previous section, we have listed the available sensors on the mobile phone, and in this section, our aim is to give an overview about the activities that can be recognized using these sensors identified in the state-of-the-art activity recognition systems on mobile phones.

When we look at the early works on mobile phone sensing for activity recognition, they address coarse-grained activities associated with location information, such as staying at home or being at the office. However, these inferences do not tell much about the exact activity performed by the user. For instance, being at the office is not equivalent to

working [23], or staying at home does not tell us whether the user is watching TV or having lunch. With newer sensors available on the phones, we can go beyond using coarse locations using substitutes of the activities. For instance, by using the information from an accelerometer, identifying that the user is sitting, from a microphone, mentioning that there is a conversation going on, and from the Bluetooth sensor, that the user's office contacts are around, we can provide a more detailed recognition process and conclude that the user is sitting at a meeting in the office.

Besides the work on associating the user's location with an activity, some other early work focused on associating the user's movement with an activity. For instance, by using the fluctuations in GSM signals, it is possible to infer whether a user is in a vehicle or is in stationary state with around 80 % accuracy [92]. Compared to location-based activity recognition, this provides a better fine-grained activity recognition, but for instance, it cannot distinguish the similar activities, such that a user is running or cycling. However, using accelerometers besides the fluctuations in wireless signals can provide us a better performance in accurate activity recognition.

Location- and motion-associated activity recognition are the two dominating types of activity recognition using mobile phones. Besides these, recent applications consider using mobile phones for more complex activities, for instance, in the field of sports: outdoor bicycling, soccer playing, lying, nordic walking, rowing with the rowing machine, running, sitting, standing, and walking using accelerometers in [46], or for daily activities such as shopping, using a computer, sleeping, going to work, going back home, working, and having lunch, dinner, or breakfast in [22]. Some recent applications also consider to use mobile phones for detecting dangerous situations such as falls [26, 98].

In Table 1, we present sample types of activities that are inferred in state-of-the-art activity recognition systems on mobile phones classified into six different categories according to their objectives.

2.3 Application Domains for Activity Recognition: Health, Well-being, and Lifestyle Change

Activity recognition using mobile phones have been used or have the potential to be used for various application domains. A detailed list of applications of mobile activity recognition was presented in [57]. They classified the application domains into three categories: (a) applications for end users, such as fitness tracking, health monitoring, and fall detection; (b) applications for third parties, such as targeted advertising, research platforms for data collection, and corporate management; and (c) applications for crowds and groups, such as activity-based social

Table 1 Types of activities studied in the literature

Class	Activity types
Locomotion	Walking, running, sitting, standing, still, lying
Mode of transportation	Biking, traveling with a vehicle, riding a bus, driving
Exercise	Outdoor bicycling, soccer playing, biking on a fitness bike
Health related activities	Falls, rehabilitation activities, following routines
Daily activities	Shopping, using computer, sleeping, going to work, going back home, working, lunch, dinner, breakfast, in a conversation, attending a meeting, using an ATM
Usage of the phone	Text messaging, making a call, browsing the web, composing an email, using an app

networks, and place and event detection. In this section, we summarize particularly the applications for health, well-being, and lifestyle changes that can benefit from the mobile activity recognition research.

The high correlation between the level of physical activity and the level of well-being is one of the key enablers of using mobile phone-based activity recognition in *health-care* applications. Common diseases such as obesity or hypertension are all linked to physical inactivity. In the current practice, the patients are asked to keep a diary about their physical activities throughout the day. The success of the diary approach depends on the user's willingness to keep everything written. However, an automatic activity recognition system based on mobile phone sensing can offer a more reliable and flexible solution. Keeping precise information on the user's activities can potentially improve the treatment of a disease.

Activity recognition on mobile phones can also help to follow the daily habits and routines of users, especially the elderly. Deviations from routines can easily be identified in such applications, and this can assist the doctors or caregivers to diagnose conditions that would not be observed during routine examinations [59].

Similarly, mobile activity recognition can be used for the rehabilitation of diseases. For instance, an activity recognition system can detect if a user is correctly doing the exercises recommended by a physician [14]. Another field within the health-care domain would be to recognize the relationship between a user's physical activity level and mental condition. Especially to the elderly experiencing dementia risk or Alzheimer who show inconsistencies in

their daily routines. An automatic activity recognition system, summarizing a user's daily routine would be beneficial to keep the progress of her mental condition and status of the disease. One challenge about this application domain, especially considering to work with the elderly, is that the elderly may experience difficulties in interacting with the smart phone interface due to their limited experience with technology and experienced impairments at different levels [93].

Well-being and *fitness monitoring* are also typical applications targeted in the mobile activity recognition studies. The mobile phone can act as a pedometer in monitoring the step count and can easily track distances traveled and calories burned [9, 70]. Additionally, using persuasive techniques, mobile phones can interact with the users to change their behavior and lifestyles in being more active [25, 81].

Ambient-assisted living is another application domain within the health care that can benefit from activity recognition systems. Assistance for people with cognitive disorders or people with chronic conditions can be provided, and their daily physical activities and routines can be monitored with a mobile activity recognition system. Mobile phone-based fall detection is another application domain exploited by the researchers recently [26, 98]. Particularly in [98], we focused on detecting falls which is considered as a major obstacle to independent living not only for the elderly but also to patients with neurodegenerative diseases, such as epilepsy. When a fall is detected especially outdoors, the proposed system also supports online location identification using GPS available on the smart phones.

3 Process of Activity Recognition

The activity recognition process can be summarized as determining a target set of activities, collecting sensor readings, and assigning sensor readings to the appropriate activities. In other words, it is the process of how to interpret the raw sensor data to classify a set of activities. Many of the activity recognition studies, not necessarily in the field of mobile phone sensing, focus on the use of statistical machine learning techniques[2] to infer information about the user activities from raw data. The learning phase can be supervised or unsupervised. Supervised techniques rely on labeled, i.e., associated with a specific class or activity, sample observations to build classifiers, whereas unsupervised techniques do not rely on labeled data. Since an activity recognition system returns a label of an activity,

---

[2]Although activity recognition can be performed using rule-based inference or using unsupervised techniques, it can be very challenging to discriminate activities in this context [54].

such as running and walking, usually they follow supervised approaches or semi-supervised approaches where a part of the training data can be unlabeled. However, in the literature, only [58] focuses on using semi-supervised learning approaches for activity recognition on mobile phones, whereas other studies utilize supervised learning techniques.

Supervised learning methods are composed of two main phases: training and classification, i.e., testing. In the training phase, machine learning approaches utilize a given set of examples or observations, called the "training set," to discover patterns from the sensor readings. These examples or observations should be associated with a specific class of activity or in other words should be "labeled" to learn from these instances. Labeling the data in the training phase is usually a tedious and complex process. Either the user should label each activity performed, such as keeping a diary or using an automatic voice recognition system to record each activity performed, or the activities of the user should be recorded with a video camera, and the activities are automatically labeled by the system. After the collection of labeled data, usually the preprocessing (noise removal and representation of raw data) and feature extraction (abstractions of raw data to represent main characteristics) steps are followed. The details of preprocessing and feature extraction are explained in Section 3.1. After these steps, training models are built, and training parameters are calculated according to the used machine learning technique. In the following sections, we summarize the main steps in the activity recognition process for the classification phase utilizing machine learning approaches and discuss the metrics used to evaluate the performance of a classification technique.

3.1 Activity Classification Steps

As discussed in [13, 79], after the sensor data are collected, the main steps of activity recognition include (a) preprocessing of sensor data, (b) segmentation, (c) feature extraction, (d) optionally dimension reduction, and (e) classification. Figure 2 shows the typical steps of activity recognition.

The *preprocessing* step contains noise removal and representation of raw data. The *segmentation* phase is usually applied to continuous stream of sensor data to divide the signal into smaller time segments since retrieving useful information from a continuous stream of data is a difficult problem. For this purpose, different segmentation methods can be applied to time-series data which enhance the signal behavior and enable us to gather useful information from continuous stream of data. The *feature extraction* phase includes the generation of abstractions that accurately characterize the sensor data. In other

words, large input sensor data are reduced to a smaller set of features, called the feature vector, that represents the original data in the best way. *Dimensionality reduction* phase can be applied to remove the irrelevant features to decrease the computational effort and memory requirements in the classification process. The aim of dimensionality reduction is to reduce the computational complexity and increase the performance of the activity recognition process. After the previous steps, the collected data can be used directly in the classification step. Finally, the *classification* phase includes mapping the sensor data (i.e., the extracted feature set) to a set of activities. The classification technique may involve a simple thresholding scheme or a machine learning scheme based on pattern recognition or neural networks. As we will elaborate in Section 5.4, common pattern recognition algorithms include decision tables, decision trees, hidden Markov models (HMM), Gaussian mixture models, and support vector machines. The reader can refer to [79] for the details and comparisons of classification models used in activity recognition research.

To wrap up, although the classification phase and algorithms give the final decision about recognizing an activity, each phase is equally important. Representing raw data without loosing the useful information in the preprocessing phase [13], efficient segmentation of continuous signals, and extracting the best features that characterize raw signals are all the key steps in delivering activity recognition results of high performance.

3.2 Performance Parameters

Figure 3 presents the possible workload and system parameters typically used in the performance evaluation of activity recognition on mobile phones. In this figure, activities and different users can be grouped as workload parameters, whereas sampling rate, phone models, classifiers, features, and segmentation type (window size) can be defined as system parameters.

- *Activity:* It is the target action being performed during tests to be recognized. Selected activity sets are also important in terms of the accuracy of the system.
- *Users:* People may perform different activities differently or can perform multiple tasks at the same time which can affect the activity recognition performance negatively.
- *Phone model and hardware:* Device model is highly important in terms of embedded sensors and computational capabilities. Sensor hardware changes according to the model and its manufacturer which may directly affect the performance as different sensors may have different accuracy and noise characteristics.

Fig. 2 Typical steps of activity recognition
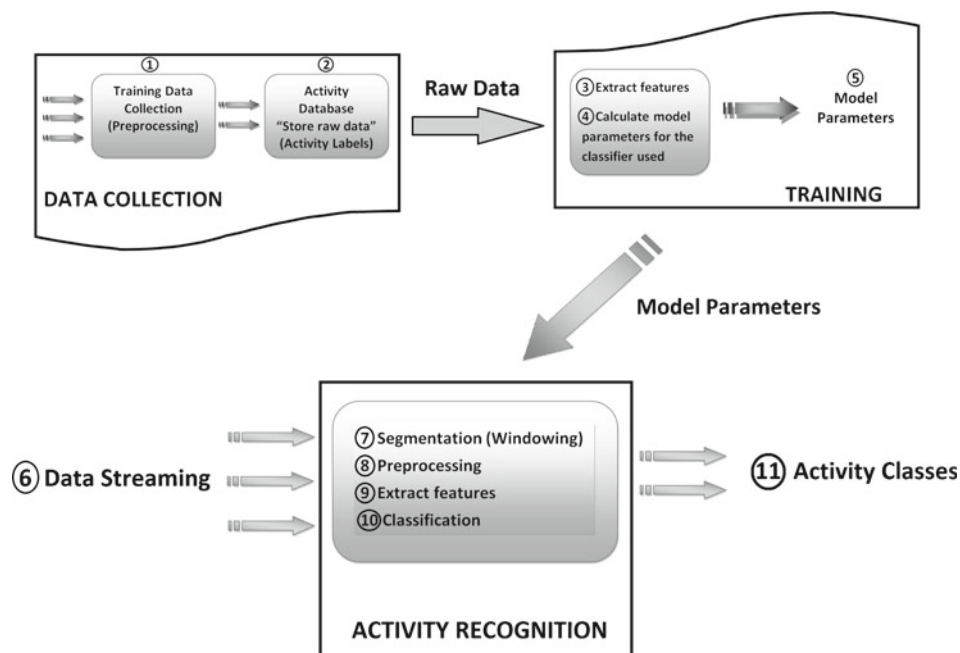
Additionally, computational power, memory, and storage capabilities of the device should be sufficient to handle the selected system parameters and test cases appropriately.

- *Sensing Modalities:* Related to the device model, sensing modalities are also important since not all the phones have the same set of sensors, and the performance of the activity recognition process may differ according to the set of sensors used. For instance, while only an accelerometer can be used to detect motion, the performance can be enhanced with utilizing more sensors such as the GPS [85].
- *Classifier:* Classification is the key step in the activity recognition process. Type of selected classifiers play an important role on the system performance.
- *Features:* Features are the signatures of the activities which have an important impact to identify the activities and affects the results directly.

- *Sampling Rate:* It is the rate at which the data are gathered. It affects the capability of the accelerometer to capture the necessary information for target activities.
- *Segmentation type (Window size):* It is the duration in which only data are collected without performing any classification. Each human activity has a pattern except the stationary activities. Because of this reason, the collected data during a window play an important role to identify activities.

3.3 Performance Measures

In the testing (classification) phase of an activity recognition system, the output classes should be compared with the ground truth, i.e., what the user was actually doing in order to evaluate the success of a classification scheme. In most of

Fig. 3 Workload and system parameters
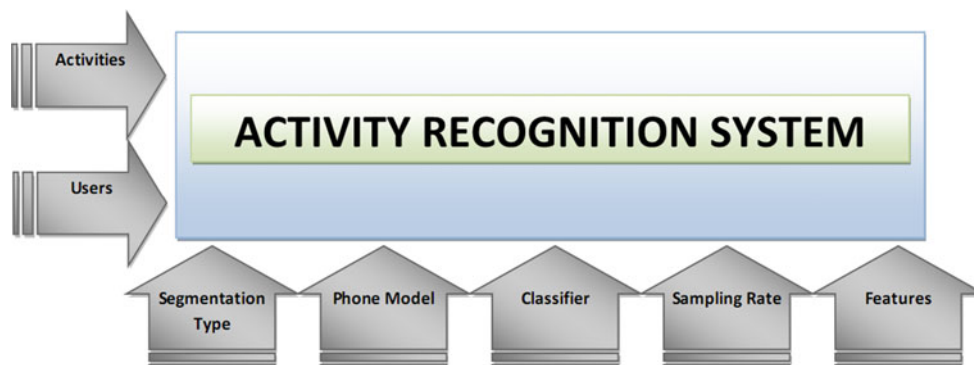
the studies, cross validation is employed using a large part of the collected data for training where the ground truth is available from the labels associated with the raw data. In some studies, training and test data may be split.

Although different studies may use different performance measures, the research community working on activity recognition is adopting similar performance measures: accuracy, precision, recall, confusion matrices, and F-measure [18].

Accuracy and precision are the mostly adopted performance measures in the literature, as outlined in Table 2. In [43], the efficiency of performance measures for activity recognition is discussed. Since we usually deal with unbalanced datasets, such that some classes may appear much more frequent than other classes, it is argued in [43] that, since the conventional use of precision and recall assumes a two-class problem (i.e., a positive class and a negative class), instead of using the precision and recall values for all activities as a metric, the average precision and recall over all activities should be computed.

4 Challenges of Activity Recognition with Mobile Phones

In this section, we first investigate the specific challenges that are only linked with activity recognition on the mobile phone environment, then the challenges related to persuasion and lifestyle change and briefly mention the general challenges of activity recognition since more detailed challenges have been discussed thoroughly in previous work [13].

4.1 Continuous Sensing

Continuous sensing, i.e, continuous sampling of the sensors on the mobile device, is a fundamental requirement for activity recognition applications, but it is also a fundamental challenge considering the battery limitation of the mobile devices. For instance, in [94], it was shown that a fully charged Nokia N95 device can support telephone conversations more than 10 h, whereas it can function only around 6 h while the GPS is turned on without taking into consideration whether it is taking samples or not. While supporting continuous sensing applications, the user experience should not be disrupted, such that the user should be able to use the phone for making calls, sending SMS's, taking pictures, or browsing on the web. In this regard, energy-efficient sensing mechanisms are required where the sensors can be duty-cycled for energy efficiency or turned off when not required to take samples. Studies that tackle the continuous sensing problem propose different approaches [24, 41, 80, 94, 96]. For instance, Wang et al. [94] propose two techniques for

reducing the battery consumption. The first one is to turn off unused sensors automatically according to user states. For this purpose, an XML-style state descriptor is taken as input and used by the sensor assignment functional block. Accordingly, selected sensors are sampled during a specific state, whereas differently selected sensors are tracking any possible user state transitions. The second technique simply uses sensor duty cycling instead of continuously sampling the sensors. Although duty cycling the sensors increases the recognition latency, using these techniques, they show that the battery lifetime can be improved by over 75 % compared to an existing application, Cenceme [66].

In [60], a continuous sensing engine, named Jigsaw, for mobile phone applications was proposed. Jigsaw sensing engine automatically adapts the GPS sampling according to the mobility mode of the user by the classification of user activities in real time using the accelerometer. Based on the user's mobility, it switches the GPS on/off with the objective of minimizing the expected localization error. Using a similar approach, an adaptive sensor sampling scheme is proposed in [24], for balancing energy–accuracy–latency trade-offs. The scheme is based on linear reward-inaction learning. According to the events observed (interesting or not), the sampling rate is adjusted, such that sensors are sampled at a high rate if the observed event is interesting (some audible data for the microphone sensor) and at a if there is no event of interest (silent event detected by the microphone). Different from these studies, in [41], energy efficiency issues are considered for multiple concurrently running applications. An engine named, Symphoney, is introduced effectively to coordinate the resource use of concurrent contending applications as well as to maximize their utilities even under severe resource contention.

4.2 Running Classifiers on Mobile Phones

As mentioned, algorithms used in the classification of activities originate from statistical machine learning techniques. However, a trendy algorithm [19] in machine learning research may not exhibit a superior performance in the field of activity recognition, especially on the mobile phone platform with limited resources, considering the limited processing power and the battery. Moreover, when we look at the literature on activity recognition using inertial sensors, we observe that most of the studies first collect sensory data and apply classification algorithms off-line on the collected data, using a large part of the collected data for training. It is clear that with the larger amount of overlap between the training data and the testing data, such as using cross validation, better recognition results will be achieved (unless the overfitting problem occurs). Off-line processing certainly exploits this advantage.

Off-line processing can be used for applications where online recognition is not necessary. For instance, if we are interested in following the daily routine of a person, such as in [97], the sensors can collect the data during a day; the data can be uploaded to a server at the end of the day and can be processed off-line for classification purposes. However, for applications such as a fitness coach where the user is given a program with a set of activities, their duration, and sequence, we might be interested in what the user is currently doing and/or if he is performing the activities with a correct sequence [97]. Therefore, online recognition of activities becomes important, especially for real-world personal fitness and well-being applications running on smart phones to provide the context of the users. Figure 4 presents a comparison of off-line/online training and classification.

Although smart phones continuously evolve in terms of computation, memory and storage capabilities, they are still resource-limited devices, and running a resource-intensive classifier may not be possible, such as when classifying audio data [51]. However, early examples of activity recognition applications on mobile phones show that classifiers such as decision trees, the minimum distance classifier, k-nearest neighbor (KNN) classifiers can run on mobile phones while providing good accuracy rates [48, 85]. Considering the resource-intensive classifiers, one approach adopted in the literature is to rely on backend servers with the uploading of sensed data to a server and benefit from their computational resources and download the results of the inferences. However, with this sort of computation, it is not possible to support real-time applications.

4.3 Phone Context Problem

One of the problems associated with mobile phone sensing is the *phone context problem* as identified in [51]. The phone context problem occurs when the phone is carried at an inappropriate position relative to the event being sensed. For instance, if the application wants to take a sample from a light sensor when the phone is located in the pocket or in the bag, the phone context problem is encountered. Especially with accelerometer-based activity recognition, the location where the phone is carried, such as in the pocket or in the bag, impacts the classification performance [51]. In most of the studies using inertial sensors, i.e., the accelerometer, the phone was restricted to be carried in a particular location by the users as we elaborate in Section 5. Recently, in two studies, the phone context problem has been investigated [67, 73]. In [67], the authors develop a system for automatic phone context discovery. In their three-level inference system, the first level is responsible for an inference from the data collected by individual sensors. However, the result of this inference may not be conclusive since individual sensor readings may not identify the context correctly, for instance, the conclusion of being in a dark environment made by the camera sensor may not be correct if the camera is covered by the user's hand. In the second step, multisensor inference is performed, such that the outputs of individual sensor inferences are combined. Finally in the third level, temporal smoothing and a HMM is applied for a final decision. The system is tested with using only the microphone sensor and for the "in the pocket" and "out of the pocket" cases and revealed a performance with around 80 % accuracy in detecting the position of the

Fig. 4 Off-line/online training and classification

phone. In a very recent study [73], Park et al. investigated the phone context problem, which was defined as device pose classification in the paper, using kernel-based estimation methods. Instead of using the built-in sensors available on a smart phone, they used Nokia SensorBox (connected via Bluetooth to the phone) including a consumer-grade accelerometer, gyroscope, magnetometer, thermometer, and barometer. However, only the accelerometer was utilized in the experiments. They use of support vector machines (SVM) and decision tree classifiers utilizing features of discrete Fourier transform (DFT) of the horizontal and vertical components of the acceleration signal as well as tilt features derived from the gravity vector that could achieve more than 98 % accuracy for off-line classification of poses bag, ear, hand, and pocket. The authors also evaluated the system performance with online recognition tests for the walking activity, and overall, the classifier predicted true device pose quite well.

Although the phone context problem has been attempted to be solved by very recent studies [67, 73], there still remains open issues such as the phone context identification in real time including different poses and for different activities.

4.4 Training Burden

Another challenge is about running the training phase of the classifiers. Even though a proposed system does online classification, the training phase can be handled with off-line processing. Usually, proper training models are being created off-line, so that these static models can be used in an online classification phase [88]. The off-line training phase is not an easy task, and it is essential for an activity recognition system since such systems require large, well-defined, and proper training sets to create appropriate models. These training sets are collected over a long period of time (couple of days, weeks, or even years depending on the related work) with dedicated and a sufficient number of test subjects as presented in [45]. Additionally, collected datasets can be too large to be processed online on other devices like smart phones because of their limited capabilities. Considering these challenges, research on human activity recognition systems explore the ways for online training.

In order to develop a real-world application where the user installs the application and does not require to deal with the burden of the training phase, such that the application can be pretrained or trained quickly without the requirement of a long training phase, and which is ready-to-use, the research question is "Can we recognize the activities with only a limited set of training data and even without any training data in a user-independent way?" This question has been partially addressed in [65] aiming user-independent or limited training of proposed systems. A system called Darwin

has been proposed to decrease the training burden on users and improve the user experience on smart phone-based on model pooling which is simply sharing and exploiting classification models which are already built by other phones and by other users. Using model pooling, there is no need to generate training models from scratch. In [47], we also evaluated the performance of classifiers with using training data from other users where the user's own training data were excluded. The results were found to be promising—giving an insight such that the training data coming from two distinct subjects with the same physical features and using the same phone model can be used for each other, which is an important hint to be able to create user-independent training datasets, as also mentioned in [65].

User-independent training was also addressed in [53], considering diverse user populations in large-scale popular mobile applications and the burden of training. An approach named "community similarity networks" (CSN) was proposed to incorporate interpersonal similarity measurements into the classifier training process. Three different similarity metrics which are physical, lifestyle/behavioral, and purely sensor data driven were used, and it was shown that the CSN approach outperforms existing approaches, such as the single model and isolated model, in classifier training considering the population diversity.

Another important work in this area is presented in [76]. In this work, the authors emphasize the known mistakes and difficulties during the labeling of the training data. It is known that collecting consistent and reliable data is a very difficult task since some contexts may be marked by users with wrong labels, which creates a necessity for instructing the users before performing any training. In these cases, although the labeled data are unreliable, they still contain valuable information. For this purpose, they propose *community-guided learning* which is a framework that trains existing classifiers with unreliably labeled data submitted by different users.

Another challenge about the training phase is labeling the activities, which also appears as a burden on the users. As mentioned, the user may be asked to keep a diary, or using an automatic voice recognition system and a video recording system, the activities can be automatically labeled by the system. In this sense, the person-independent classification systems are expected to overcome this challenge as discussed in the previous paragraph.

4.5 Phones not Always Carried

Although people consider the mobile phone as essential when leaving home besides a key and a wallet [11], the phone is not always carried. For example, people often leave the phone on the desk when at work or do not carry it around when at home. This is not, for instance, the case with

some other dedicated wearable sensors that are attached on the body. However, wearable sensors are considered to be more obtrusive than using mobile phones, and this reflects the trade-off between these monitoring modalities. Systems utilizing different sensing modalities and their integration could be a topic of interest for future research.

4.6 Persuading Users

Especially in persuasive applications for a behavior or lifestyle change, the success of the application depends on the willingness of the user to engage. Persuasion techniques such as self-monitoring, social influence, and fun interaction are often used in the design of such applications [90]. Virtual companions [12], games [78], and social networks [99] are commonly used to interact with the users in mobile applications. However, understanding which types of methods and metaphors are the most effective for various applications and people is still under investigation. As also mentioned in [51], metaphors to motivate users to use the application for persuasive computing should be investigated in more detail, possibly by interdisciplinary research teams including psychologists.

4.7 Human Behavior

People can perform multiple tasks at the same time which can affect the activity recognition process negatively. Additionally, various continuous sequences of performing tasks and their periodic variations may result in incorrect classifications. Because of these reasons, accuracy and reliability of sensor data play an important role in the activity recognition.

5 State of the Art on Activity Recognition on Mobile Phones

As we mentioned in Section 2.2, location- and motion-associated activity recognition are the two dominating types of activity recognition using mobile phones. In this section, we first review the studies that focus on location-associated activity recognition. Next, we focus on the studies based on motion-based activity recognition. Finally, we review the studies that consider different types of activity recognition other than motion and location. Figure 5 shows the types of activity recognition on mobile phones classified according to their objectives.

5.1 Location-driven Activity Recognition

Before the release of smart phones equipped with a rich set of sensors, early examples of activity recognition on

mobile phones focused on the use of location information to detect the user activities. Location-driven activity recognition aims to recognize activities associated to certain places [23]. In the Reality Mining Project [30], three activities "home, work, elsewhere" were targeted using cell tower and Bluetooth data. An HMM model, conditioned on the hour of the day and day of the week, was built, and the associated activities were recognized with 95 % accuracy. Similarly in [36], semantic content associated with locations were used to infer user activities, such as in shops, restaurants, recreation, government offices, schools, and entertainment places. Location data were collected with GPS. A very recent work that introduces a generic service for indoor and outdoor detection, named IODetector, was presented in [100]. IODetector utilizes light sensors, magnetism sensors, and cell tower signals to detect indoor/outdoor switching. Indoor/outdoor knowledge is proposed as an important information for activity recognition systems leading to more accurate recognition.

The main drawback of using only location-driven activity recognition is the inaccurate inferences. For instance, being at home does not mean eating or sleeping. Moreover, activities inferred with places are usually not of interest for personal health or wellness systems. However, if location information is used with other sensory data, for instance, using the microphone or accelerometer, it can help to improve the activity recognition results.

Another related domain in location-driven activity recognition is indoor positioning and tracking systems. For instance, detecting patterns of wandering and their locations as well as detecting hazardous situations typically for demented patients, such as closeness to a window, are also important factors in monitoring the well-being of patients especially those with cognitive disorders. Due to the lack of reliable GPS signals indoors, in most of the studies, the use of Wi-Fi beacons and inertial sensors are investigated for indoor positioning. The main challenges of indoor tracking systems are that the position of the phone, such as in a hand or pocket, bouncing of the phone when it is in a pocket, and receiving a phone call, may affect the sensor readings and make the tracking challenging [56]. Although not well linked to personal health applications, examples of indoor positioning systems can be found in [56, 74, 89], and they present a potential to be utilized in the health and well-being domain.

5.2 Motion-based Activity Recognition

Motion-based activity recognition systems mostly utilize inertial sensors, radio sensors (cellular, Wi-Fi), or other sensors such as the GPS for motion recognition. In this section, we first review the systems that utilize transceiver interfaces

Fig. 5 Types of activity recognition systems on mobile phones
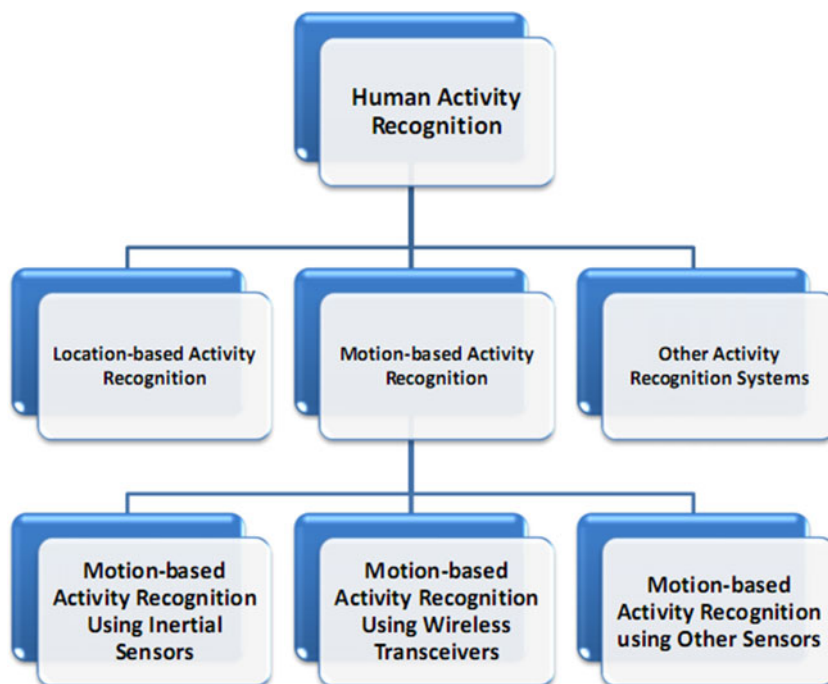
on the phones and, next, summarize the systems that utilize inertial sensors, and finally look at the studies that utilize the combination of these sensors or use other types of sensors such as the GPS.

5.2.1 Motion-based Activity Recognition using Transceiver Interfaces

In [92], activities of walking, driving, and staying at the same place (dwelling) were identified using only the GSM traces. Additionally, a GSM-based step counter was also proposed. The principle behind the mobility mode inference is that radio signals observed from stationary sources are fixed in time but variable in space. By observing a series of GSM signals from a set of stable towers and calculating the Euclidean distance between consecutive GSM measurements, mode of transportation is inferred since stationary, slow and fast walking and driving show different distance values. A boosted logistic regression technique was used for the classification of activities. By measuring the walking periods and assuming an appropriate step rate, the user's daily step count was also calculated. The performance of the system was evaluated with three users for 1 month using the Audiovox SMT 5600 mobile platform. The overall accuracy for the transport mode recognition was calculated as 85 %, and the step counter reasonably estimated the number of steps calculated by commercial pedometers. Although a specific application was not targeted in the paper, elderly care and monitoring the elderly's

wellness through mobile phones were mentioned as suitable applications regarding the personal health and well-being systems.

Shakra [10] is an application for tracking and sharing daily activity levels with mobile phones. Shakra also utilizes GSM signal traces to detect the activities of walking, driving, and in stationary state. The system targets to track the daily exercise activities of people and users are supported to share and compare their activity levels with others for motivating fitness and moderate activity. For the classification, both an artificial neural network and an HMM were used. The application was evaluated with three groups of users with different activity levels, ranging between inactive and highly active, to find out whether the application increased the awareness of the users for their activity levels and persuaded them to be more active. Overall, the HMM performed slightly better than the artificial neural network classification revealing an accuracy of 82 %. All of the users were reported to be responding positively about the system to increase their activity levels. However, as the authors mention, it is not possible to claim that the users would remain motivated about the system, and they mention that this study shows that Shakra is usable, and, if further, developed it can act as an effective health promotion tool.

In [49], Wi-Fi signals were used to infer whether the user is moving or stationary. Whereas in [71], a hybrid approach utilizing both Wi-Fi and GSM signals were used to infer the transportation mode.

5.2.2 Motion-based Activity Recognition using only Inertial Sensors

Accelerometers on the mobile phone were initially included to enhance user experience by automatically changing the orientation of the display according to the orientation of the mobile phone held by the user. However, as widely adopted by the activity recognition research in other areas, i.e. activity recognition using body-worn sensors, accelerometers embedded in the mobile phone also have the potential to recognize the user's motion and activities.

In [88], a classification algorithm is aimed to run on the iPhone platform. For this purpose, three iPhone applications are developed called as iLog, iModel, and iClassify. iLog is used for data collection of different activities in real time. iModel is a desktop tool for learning and testing models. Data saved by iLog can be imported into iModel which is a Java application built on the Weka machine learning toolkit [2]. By using iModel, the labeled data can be used to test an existing model or to learn a new model. Lastly, iClassify can be used to classify walking, jogging, bicycling on a stationary bike, and sitting activities in real time in contrast to many studies that focus on off-line classification of the data. In addition to iPhone three-axis accelerometer sensor, Nike+iPod Sports Kit is used to collect data which imports the collected data periodically to iPhone via Bluetooth connection. iClassify can report activity classifications once per second. It performs nearly 97 % accuracy in recognizing activities.

ActiServ, a service-based recognition architecture, is presented in [16]. One of the most important features of the architecture is that it requires minimal personalization effort by the users, only 1-3 min of data collection per activity seems sufficient. The performance of the system is evaluated with 20 users, each user generating 2-3 min of data for each of the activity classes, using OpenMoko Neo Freerunner phone with extensive test scenarios. If the training data of the user evaluating the system is used with the phone located in the same position both in the training and evaluation phases, then the performance of the system is reported to be 97.3 % in terms of accuracy. If the model is trained when the phone is in different position, then the accuracy rates drop to 60 %. When the training data of the evaluation user are excluded and the data from other users (actually selected from the best matching users) is used, the accuracy rates are between 86 and 63.6 %. Although the data were collected on a mobile phone platform, ActiServ is both trained and tested off-line on backend servers. ActiServ does not detail a target application, but the activities classified and off-line processing can be used for monitoring the daily routine of a person, such as a personal activity diary for a wellness application.

Probably the most common device to promote more active life is the use of pedometers which allow for step counting [9]. Examples of pedometer applications developed for smart phones are available in application stores. Similarly, a step counter service was proposed in [70] whose performance was compared with different step counting products. The service used hill detection and threshold calculation for step identification utilizing the magnitude of acceleration and detected the orientation change to improve the performance of recognition. The performance of the server was evaluated with a single user by conducting walks of 500, 1,500, and 2,500 m, and each walk was repeated three times. According to the experimental results, the service on the average reveals a mean error of 0.5 % compared to the reference Nike+ sports kit. The performance was also compared with step counting products such as the Nokia step counter with the phone carried in different positions (pocket, belt clip, hand, backpack), and it exhibited better results.

In [17], instead of periodic activities such as walking, cycling, and car driving, transitions between physical activities, particularly between sitting and standing, are targeted. Assessment of the behavior of elderly people and pregnant workers facing inappropriate working conditions under ergonomic aspects are the target stakeholders for this work. A set of kernel functions were utilized, and cross correlations with the sampled and projected signals were calculated to detect a matching. A correlation coefficient near 1 infers a sit-to-stand transition, whereas near -1 infers a stand-to-sit transition. The performance of the system was evaluated on SonyEricsson w715 phones with 12 subjects, and 70 % average recognition rate was achieved. Although the target for the system was the detection of inappropriate working conditions or in support for the health conditions of the elderly, it was not tested under the circumstances of these applications.

In [97], instead of only focusing on the classification performance of activities, the objective is to provide a physical activity diary, with a potential application of mobile health care. A rich set of 15 different features (vertical features, horizontal features, and cross-correlation features) were extracted from raw acceleration signals, and decision tree (C4.5), naive Bayes (NB), KNN, and support vector machine learning algorithms are used as the classifiers. It was shown that the decision tree achieves the best performance among the other classifiers, with around 90 % accuracy. In order to create a physical activity diary, the system was used by a subject for several months.

In a recent study, Park et al. [73] focus on the classification of the mobile phone's position relative to the body and estimation of walking speed only using the accelerometer. Phone position classification is used to tackle with the

phone context problem as discussed in Section 4. Walking speed estimation is used for health monitoring applications. Regularized kernel methods are used for the classification together with the features of raw acceleration data, such as the discrete Fourier transform, horizontal, and vertical components. The system is implemented on the Nokia N900 platform and tested with 14 subjects. The median absolute speed error was reported to be 0.039 m/s, which amounts to 3 % of the average walking speed.

Similar to the other studies, Kwapisz et al. [50] concentrated on recognizing the common human activities such as walking, jogging, ascending stairs, descending stairs, sitting, and standing. They used inertial accelerometer sensor of Android-based smart phones. This study differs from other studies in terms of the data size collected for the training phase which is performed by 29 different subjects. The classification steps are performed with the help of the Weka toolkit by using a decision tree, logistic regression, and multilayer neural networks. They used ten-fold cross validation for all experiments. According to the results, multilayer perceptron achieved the best performance with 91.7 % accuracy rate.

As discussed in Section 2.3, activity recognition may target different application domains. In this context, examples of gait analyzers have been presented in [38, 39] for health-care and presence services. Additionally, Fontecha et al. [34] proposed a complete elderly frailty detection system by using accelerometer-enabled smart phones and clinical information records of the subjects. There are many different factors to be evaluated for frailty detection and diagnosis. In this manner, assessment of physical condition through gait and other physical exercises is one of the most important factors. This work presents a mobile system to collect elderly data based on gait and balance tests. They create instances for each subject with the combination of dispersion measures coming from accelerometer data and risk factors from patient records. By comparing different instances, they create an affinity tree which is used for frailty diagnosis.

Differently from other studies, Zhixian et al. concentrated on energy efficiency and introduced an activity-sensitive strategy ("A3R" - adaptive accelerometer based activity recognition) for continuous activity recognition in [96]. They studied the individual's locomotive activities such as standing, slow walk, sit–relax, sit, normal walk, escalator-up, escalator-down, elevator-up, elevator-down, and downstairs. The proposed system achieved an overall energy saving of 20–25 % by adapting the accelerometer sampling frequency and the classification features separately for each activity in real time.

As we have summarized, most of the algorithms focus on simple activities such as locomotion. In [29], more complex activities are targeted to be monitored, such as cleaning,

cooking, medicating, sweeping, washing hands, and watering plants, besides the simple activities, such as biking, climbing stairs, driving, lying, sitting, standing, and walking. The activity recognition system was developed on the Android platform. Weka machine learning toolkit was used to test six different classifiers: multilayer perceptron, naive Bayes, Bayesian network, decision table, best-first tree, and K-star. Although the classification accuracies for simple activities were found to be above 90 % except the naive Bayes classifier, the best accuracy achieved for complex activities was around 50 %. Simple activities could retain their high classification accuracy even when tested together with the complex activities.

Similarly in [95], locomotive microactivities are used to identify semantic activities such as cooking, working, eating, relaxing cleaning, on a break, and meeting. Using lifestyle data with five users for 152 days, the proposed approach is reported to achieve an average accuracy of 77.14 %. with a 16.37 % improvement compared to the one-tier approach. The same authors investigate the recognition of complex activities using higher order features and SVM-based fusion mechanisms in [84]. The same dataset was used, and they reported an average accuracy of 86.17 % for the same set of activities.

As we mentioned in the list of challenges in Section 4, the training phase of the algorithms using machine learning approaches creates a burden on the users. The research community is interested in building classification models where the activity recognition phase can be performed in a user-independent way without the requirement of training data collection by the user [61]. User-independent activity recognition on mobile phones has been recently addressed in [91]. In this work, the targeted activities are walking, running, cycling, driving a car, and idling (sitting/standing). In the experiments, in both the training and testing phases, the phone was carried in the pocket of the subject's trousers. Twenty-one features in total were extracted from the magnitude of acceleration, including standard deviation, mean, minimum, maximum, five different percentiles (10, 25, 50, 75, and 90), and a sum and square sum of observations above/below a certain percentile (5, 10, 25, 75, 90, and 95). In order to recognize activities, a static decision tree is first used to detect whether the user is active or inactive. If he is inactive, then the classes to be recognized are sitting/standing or driving. But if the user is active, then the classes to be recognized are walking, running, and cycling. The second stage is performance of both KNN and quadratic discriminant analysis (QDA). The performance of the classifiers is tested both with off-line and online classification, and in the off-line case, QDA revealed 95.4 % accuracy, while KNN performed with an accuracy of 94.5 %. In the online case, the accuracy achieved with QDA was 95.8 %, whereas it was 93.9 % for KNN. In the online tests, seven

subjects tested the system, and training data of three subjects were excluded to make user-independent inferences. Although the reported recognition rates are quite high, the authors report two cases where user-independent classification has not performed very well. The walking activity of one of the subjects whose training data were not used was recognized with 65.6 % accuracy when QDA was used, and cycling activity of another subject whose training data were not available was recognized with an accuracy of 76.3 % using KNN. It is reported that in both cases, cycling and walking were mixed together. Authors also report that walking is one of the most challenging activities to recognize user-independently since different subjects have different personal walking styles.

Although in [91] the classification phase was performed online on the mobile platform, together with feature extraction and segmentation phases, the training phase was performed off-line, similar to that in [66, 85]. In [47, 48], we focused on online recognition of activities on smart phones together with performing the training phase also on the mobile platform. This work is motivated by the fact that considering the challenge of training burden mentioned in Section 4, research on human activity recognition systems explore the ways for online training [31]. We also focused on the common activities targeted in the literature: sitting, standing, walking, biking, and running. We used the three different classifiers naive Bayes, clustered KNN, and decision tree using mean, maximum, minimum, and standard deviation in the feature set. These classifiers were selected considering the limited processing and storage on smart phones and since these classifiers and features were commonly used in the previous studies—using the same set makes it easier to compare our findings with the similar studies. Performance of the classifiers is tested with ten different subjects on different Android-based mobile platforms considering the most effective system parameters like the window size and the sampling rate. According to the results, the clustered KNN method exhibited a much better performance with an accuracy around 92 % excluding biking and 73.4 % accuracy with biking activity. The naive Bayes classifier performed with 48 % accuracy excluding biking activity and 32 % accuracy including biking. On the other hand, its performance is nearly the same as the decision tree classifier which performed with 86 % accuracy excluding biking and 76 % accuracy with biking. Similar to [91], we also evaluated the user independency of the system in the training phase. When we excluded the training data of the users in the classification phase and used only the training data from other users, the average accuracy dropped to 48 % with the decision tree algorithm including the biking activity. However, for the users who performed the tests on the same platform the accuracy was not affected, although they did not use their own training data. We are currently

working on this issue to better evaluate the user independency of activity recognition using these classifiers on the Android-based platforms.

5.2.3 Motion-based Activity Recognition using Other Sensors

Reddy et al. [85] proposed a different model for activity recognition. In this study, they designed, implemented, and evaluated a transportation mode classification system which runs on a mobile phone by using a three-axis accelerometer and GPS sensors. Simply, they focused on outdoor activities and classified them into the following five groups: walking, in stationary state, biking, running, and motorized transport. In fact, they first considered the use of different sensors, such as Bluetooth, Wi-Fi, and GSM cell radio. However, experimental evaluation concluded that the most dominant sensors in the performance improvement are the accelerometer and the GPS sensors. Although the GPS sensor was observed to be the most dominant one, it was not successful in identifying the biking and running activities, since they reveal similar speed patterns. To distinguish these activities from each other, they used the acceleration sensor and could achieve a recognition accuracy over 93 %.

Similarly in [87], a system called *ambulation* was proposed for monitoring the mobility patterns of using the accelerometer and GPS sensors available on mobile phones. The system targets patients who suffer mobility-affecting chronic diseases, such as Parkinson's disease. Variance, mean, and fast Fourier transform (FFT) coefficients from raw acceleration data are used as the features together with the decision tree classifier. Using acceleration, the system can identify the activities of walking, running, and in stationary state. By also using the GPS sensor, additional activities of biking and driving are also identified. The system also displays the mobility traces of people, and through these traces, one can easily detect anomalies in the change of mobility behavior.

Differently from other studies, Martin et al. [62] investigated the effects of using a different set of sensors on the overall system performance while considering different factors. In this work, they emphasize that priori information on the orientation and the placement of the device relative to user's body may enhance the results of the system in terms of accuracy. For this purpose, they also used proximity, light, and magnetometer sensor data in addition to the accelerometer sensor data. The system performance is evaluated by using lightweight classification techniques such as the naive Bayes, a decision tree, and a decision table. According to the results, a computationally low-cost decision table with best suitable feature set can achieve 88 % accuracy rate with all the sensors listed previously.

5.3 Activity Recognition Systems Utilizing Other Context Information

Besides motion and location-based activity recognition systems, there are also examples where different context information, such as ambient light, noise, and even social interactions, can be used to infer activities. In this section, we first review the studies that focus on utilizing other sensors for more comprehensive context and activity recognition and then elaborate on the studies with the objective of monitoring the social interactions and relations between users using smart phones.

In [94], a novel design framework, called EEMSS, for energy-efficient mobile sensing is proposed. A hierarchical sensor management strategy is used to recognize user states as well as to detect state transitions. User states may contain a combination of features such as motion (running, walking), location (staying at home or on a freeway), and background condition (loud or quiet) which all together describe the user's current context. The state transition system implemented on Nokia N95 can currently define the following states: walking, on a vehicle, resting, at home talking, at home entertaining, working, meeting, in the office–loud, in a quiet place, place speech, and in a loud place. The sensors used for activity recognition are accelerometer, Wi-Fi detector, GPS, and microphone. EEMSS is able to detect states with approximately 92.56 % accuracy by processing testing data off-line and improves the battery lifetime by over 75 % compared to an existing application, Cenceme [66].

In [66], authors present design, implementation, and evaluation of a social networking application, called Cenceme. Basically, users are able to share their contextual information through social platforms like Facebook and MySpace using this application. Cenceme benefits from off-line computational powers of backend servers for training whereas it additionally performs online classification. In this study, the authors have focused on sitting, standing, walking, and running activities as well as on audio info of the environment. For this purpose, they used accelerometer, Bluetooth, audio sensors, and GPS which are embedded in Nokia N95. The classifiers used in this study can be grouped as audio and activity classifiers. The audio classifier benefits from DFT, and its feature vector consists of the mean and the standard deviation of the DFT power. On the other hand, feature vectors in the activity classifier are the mean, the standard deviation, and the number of peaks per unit time. Simple three-level decision tree performs classifications online on the Nokia platform for identifying the activities of walking, running, standing, and sitting states. Additionally, backend classification derives contextual information from collected data. The performance of the proposed system is tested with 22 subjects over 3 weeks on the Nokia N95 platform.

In [52], the design of an automated well-being application for the Android smart phones, BeWell, was introduced. BeWell monitors sleep, physical activity, and social interactions and gives feedback to promote better health. GPS and accelerometer are used to monitor a person's physical activities, such as driving, in stationary state, running, and walking, and microphone is used to recognize social interactions by identifying voicing and non-voicing states. Sleep durations are approximated by measuring phone usage patterns, such as phone recharging, movement, and ambient sound. BeWell also provides an ambient well-being interface to give continuous feedback about a users current well-being state. The system was tested with a single user for a week.

Sleeping patterns and sleeping hygiene can have an important effect on various health conditions, such as affective disorders, hypertension, and heart disease [8]. Besides the research activities on monitoring sleeping behavior [7, 52, 72], there are also applications available in the market [3, 4]. Mainly, the accelerometer that detect the movements of the user and microphone that record ambient noise and snoring are utilized to monitor sleeping behavior. In some of the systems [4], the user is given feedback and woken up in the lightest sleep phase.

Social interactions and relations between users are also considered as important factors in well-being monitoring [63, 64]. As we mentioned, in the "Reality Mining Project" [30], co-location of Bluetooth radios on the mobile phone were used to infer social interactions between phone users. Gaussian mixture model (GMM) was used to detect patterns in proximity between users and identify the type of relationship, such as work groups. The data were collected from 100 users for over the course of an academic year, summing up to 450,000 h of information. In a similar study [82], the SociableSense platform was introduced to monitor user interactions in office environments by providing users a quantitative measure about their sociability and that of colleagues. Colocation patterns, using the Bluetooth sensors, and interaction patterns, using the microphone, were used to measure the sociability of users. Additionally, the accelerometer was also utilized to detect the moving and stationary states of the users. The system was tested with ten users for 10 days in an office environment.

Emotionsense [83] is another platform that monitors the social interactions among users using similar techniques as in [82] but additionally recognizes the emotional states of the users with the microphone, such as happy, sad, and neutral. GMM classifiers were used in emotion recognition, and the system was tested with 18 users for 10 days.

5.4 Taxonomy

In this section, we provide a taxonomy of the reviewed examples of activity recognition systems using mobile phones according to different metrics ranging from the types of sensors used to the number of subjects included in the experiments. Table 2 summarizes the general aspects of the studies that are discussed in this section, ordered by the type of activity recognition, followed in the system, i.e., motion-based and location-based. Looking at the dates of publications, most of the studies have been presented quite recently—in the last couple of years. The most commonly used sensor is the acceleration sensor. Wi-Fi, GPS, or any other sensors are added to improve the sensing power and accuracy of the results. *Sitting*, *standing*, *walking*, *running*, *driving*, and *bicycling* are the common activities being targeted to be recognized in the applications. There are also other studies which target to recognize contextual information, such as location and environmental audio data, of the users as in [94].

From the perspective of personal health and well-being systems, most of the studies target applications such as fitness applications or tracking of daily activities. In earlier studies, the classification algorithms are implemented on Nokia N95 which is one of the first mobile phones with sensing capabilities. In the recent studies, Android phones are commonly utilized. Very diverse set of classifiers (decision tree seems to be the most common one), and features have been used in the papers with different number of test subjects that make it difficult to compare the performance of the studies even those performed with the same set of activities. As we discuss in Section 6, the collection of open datasets is required for benchmarking. Additionally, the number of test subjects in most of the studies is quite limited, even in some of them, less than ten test subjects were included. This makes it difficult to generalize the findings of these studies. Accuracies differ between approximately 80 to 97 % depending on the set of activities used and the processing techniques.

6 Open Issues and Future Research Directions

In this paper, we have provided a review of existing activity recognition systems on mobile phones. As can be seen from the date of publications in the list of references, it has become a quite hot topic in the recent years with the release of smart phones equipped with a variety set of sensors. However, there are still significant open issues, such as the recognition of composite activities rather than locomotion activities, which require further research. Moreover, the current performance results can be improved and extended.

In this section, we present a list of some possible future research directions.

- **From Locomotion Activities to Complex Activities:** Most of the studies infer locomotion activities, such as walking and running. However, the link between these basic activities and more complex activities and the context information of the user is weak. For instance, it is rather straightforward to detect if the user is running, but inferring if the user is running away from danger or jogging in a park is different [40]. Although some initial work [29, 95] attempted to address this issue, recognition of more complex activities and mapping these activities to the application domains where this activity information can be useful should further be explored.

- **Fusing Sensor Information for More Accurate Context Recognition:** The most common sensor used in the presented studies is the acceleration sensor, whereas in some of the studies, the GPS and microphone also accompany the accelerometer. Which sensors should be used together for which types of target applications for better context recognition should be explored. Besides the embedded sensors in the mobile platform, external sensors, such as the ones measuring physiological information attached to the user body, can also be utilized, or the ambient sensors, such as the ones deployed in smart homes for user behavior recognition [44], can be used together with the mobile phone sensors for a complete user behavior recognition.

- **Continuous sensing and inference duty cycling:** As mentioned in Section 4, continuous sensing is both a requirement and also a challenge for mobile activity recognition systems. The sensor sampling should be adapted while considering the battery lifetime and inference accuracy trade-offs. Although some initial work has been presented [24, 41, 60], intelligent mechanisms for duty cycling of sensor sampling are required to be further investigated so that while supporting continuous sensing applications, the user experience is not disrupted, and battery lifetime is not seriously affected.

- **User-Independent Systems:** The data collection and labeling in the training phase of the supervised machine learning algorithms are challenging tasks and may decrease the adoption of the activity recognition systems. Hence, user-independent systems with a high recognition rate should further be explored for the success of these systems. Moreover, the use of unsupervised learning techniques can also be investigated although it is a challenging task.

- **Group Activity/Behavior Understanding:** Most of the studies propose to recognize the activities of

Table 2 Taxonomy of activity recognition systems on mobile phones

Reference	Sensors	Activities	Application	Platform	Classification	Features	No. of subjects	Performance result
Eagle et al. [30]	Cell tower and bluetooth	Home, work, elsewhere	Social interactions between users	MIDP2-enabled mobile phones, Nokia 6600	HMM	Static Bluetooth device ID, staying in the same location, charging, missed phone calls, text messages, alarms etc.	100	95 % accuracy
Hariharan et al. [36]	GPS	Shopping, restaurants, recreation, government offices, schools, entertainment	User patterns	Web-connected mobile computer and GPS	Bayesian network model	Timing, velocity, frank loss of signal after a slowing of velocity, web-based location resources	–	–
Sohn et al. [92]	GPS, cell tower	Walking, driving, dwelling, step counter	Elderly care and monitoring elderly's wellness	Audiovox SMT 5600	Logistic regression	Spearman rank correlation coeff., no. of common cell towers, mean Euclidean distance, variance in Euclidean dist. values, variance in signal strengths	3	85 % accuracy
Anderson et al. [10]	GSM signal traces	Walking, driving and stationary	Tracking and sharing daily activity levels	Windows smart phones, i-mate sp5s	Artificial neural network, HMM	Means and variances along the two dimensions	2,3,4 (3 groups)	82 % accuracy (HMM)
Krumm, [49]	Wi-Fi signals	Moving, stationary, location info			HMM	Wi-Fi signal strength features (e.g. variance)	10 walks	87 % accuracy
Mun et al. [71]	Wi-Fi and GSM signals	Dwelling, walking or driving	Transportation studies, urban planning, health monitoring and epidemiology	Nokia N95	Decision tree	Signal strength variance, duration of dominant Wi-Fi access point, no. of unique cell IDs, residence time in a cell footprint	–	83 % accuracy
Saponas, [88]	Accelerometer, Nike+iPod Sports Kit	Walking, jogging,	Fitness applications bicycling, sitting	iPhone	Naive Bayes	Extracted from 124 distinct features	8	97 % accuracy
Berchtold et al. [16]	Accelerometer	Standing, sitting, lying, walking, climbing stairs, cycling, stationary		OpenMoko Neo Freerunner	Fuzzy inference systems	Variance and mean values	20	97.3 % accuracy

Table 2 (continued)

Reference	Sensors	Activities	Application	Platform	Classification	Features	No. of subjects	Performance result
Mladenov et al. [70]	Accelerometer	Step identification	Middleware, creating a layer btw. client app. and raw sensor access calculation	Nokia N95	Hill detection and threshold	Utilizing the magnitude of acceleration	1	Mean error of 0.5 % compared to the reference Nike + sports kit
Bieber et al. [17]	Accelerometer	Transitions between physical activities	Assessment of inappropriate functions working conditions under ergonomic aspects	Sony Ericsson w715	Matching - kernel and cross correlations with the signals	Mean and variance	12	70 % average recognition rate
Yang et al. [97]	Accelerometer	Sitting, standing, walking, running, driving and bicycling	Physical activity diary, healthcare	Nokia N95	Decision tree, Naive Bayes, KNN, SVM	15 different features (vertical, horizontal and cross-correlation features)	12	90 % accuracy
Park et al. [73]	Accelerometer	Estimation of walking speed	Health monitoring	Nokia N900	Regularized kernel methods	Discrete Fourier transform, horizontal and vertical components of raw acceleration data	14	0.039 m/s median absolute speed error
Henpraserttae et al. [37]	Accelerometer	Sitting, lying, standing, walking, running and jumping		Accelerometer embedded mobile-phone	Rules and threshold based classification, instance-based learning	Mean and standard deviation	10	90 % accuracy
Kwapisz et al. [50]	Accelerometer	Walking, jogging, ascending stairs, descending stairs, sitting, standing	Customization of the mobile device's behavior based on activity, health care	Android-based cell phones	Decision tree, logistic regression, multilayer neural networks	Average, std. dev., avg. absolute diff., avg. resultant acceleration, time between peaks, binned distribution	29	91.7 % accuracy (MNN)
Iso et al. [39]	Accelerometer	Walking, running, going up/down stairs, and fast-walking	Health care and presence services	Three-axis accelerometer mounted on a cell phone	Gait appearance probabilities based on Bayesian theory	Wavelet components, periodograms of the best basis and the momentum of the information entropy distribution of the best basis	2	80 % accuracy

Table 2 (continued)

Reference	Sensors	Activities	Application	Platform	Classification	Features	No. of Subjects	Performance result
Fontecha et al. [34]	Accelerometer	Collecting elderly data based on gait and balance tests	Frailty detection and diagnosis	Android-based smart phones	Affinity tree	Dispersion measures such as arithmetic mean, Std. dev., amplitude, variance, absolute mean diff., acc. mean, pearson's coefficient of variation	20	80 % precision (HMM)-
Lee et al. [55]	Accelerometer	Stand, walk, run and stair up/down, shopping, taking bus, moving by walk		Android-based smart phones	Hierarchical HMM	x, y, z acceleration values	3	
Zhixian et al, [96]	Accelerometer	Standing, slow-walk, sit-relax, sit, normal-walk, escalator-up, escalator-down, elevator-up, elevator-down, downstairs	Energy efficient continuous activity recognition	Android-based smart phones	J48 adaptive decision tree classifier in the Weka toolkit	mean, magnitude, variance, covariance, energy based on FFT component, entropy based on FFT histogram	4	Energy saving of 20–25 %
Hynes et al. [38]	Accelerometer	Walking, resting	Medical diagnosis	Java ME mobile handset	Average magnitude difference function based algorithm	Average magnitude difference, normalized average magnitude difference	10	10 % reduction in power consumption
Dernbach et al. [29]	Accelerometer	Cleaning, cooking, medication, sweeping, washing hands, watering plants, biking, climbing stairs, driving, lying, sitting, standing and walking	Technology-driven assistive healthcare	Samsung Captivate	Multilayer perceptron, NB, bayesian network, decision table, best-first tree and K-star	Mean, min, max, standard deviation, zero-cross and correlation	10	90 % accuracy for simple activities, 50 % accuracy with complex actvities
Yan et al. [95]	Accelerometer	Cooking, working, eating, relaxing Cleaning, on a break, meeting	Web log mining, mobility data mining, Social network data mining	Nokia N95	Decision tree, Naive Bayes, Bayesian network, LibSVM, Adaboost	Mean, variance, two-axis correlation, FFT magnitude, FFT energy and FFT entropy	5	77.14 % accuracy

Table 2 (continued)

Reference	Sensors	Activities	Application	Platform	Classification	Features	No. of subjects	Performance result
Siirtola et al. [91]	Accelerometer	walking, running, cycling, driving a car and idling (sitting/standing)		Nokia N8, Samsung Galaxy Mini	decision tree with KNN and QDA	21 features (e.g. magnitude of acceleration, standard deviation, mean, minimum, maximum)	8	95.4 % accuracy (offline QDA), 94.5 % for KNN
Kose et al. [47, 48]	Accelerometer	Walking, running, cycling, sitting, standing		Android phones	KNN, NB, decision tree	Min, max, average, variance	10	92 % with KNN
Reddy et al. [85]	Accelerometer, GPS	Walking, stationary, biking, running, and motorized transport	Physical activity monitoring, personal impact and exposure	Nokia N95, iPhone, TMobile G1	Decision tree, KMC, NB, NN, SVM, C-HMM, DT-HMM	Magnitude of 3 axis of acc., mean variance, energy, DFT energy coefficient btw 1-10 Hz, speed	16	93 % accuracy
Ryder et al. [87]	Accelerometer, GPS	Walking, running, stationary, biking, driving	Monitoring healthcare	Android phones, Nokia N95	decision tree	variance, mean, FFT coefficients	–	–
Martin et al. [62]	Accelerometer, magnetometer, gyroscope, proximity, light	Walking at different paces (slow, normal, rush), running, sitting, standing	Wellness and pathology treatment	Google Nexus S	Naive Bayes, decision tree, decision table	Mean, variance, zero crossing rate, 75 percentile, interquartile, correlation btw. sensors, power spectrum, centroid, FFT energy, frequency-domain entropy, signal energy	16	88 % accuracy with Decision Table
Wang et al. [94]	Accelerometer, Wi-Fi detector, GPS, microphone	Walking, vehicle, resting, home talking, home entertaining, working, meeting, office loud, place quiet, place speech, place loud	Customized applications like: Ringtone adjustment placement, Medical Applications, Personal Safety Applications	Nokia N95	Decision tree	3 characteristic features that defines each of state: location, motion, background sound (e.g. FFT, magnitude of acc. data)	10	92.56 % accuracy, 75 % improvement at the battery lifetime

Table 2 (continued)

Reference	Sensors	Activities	Application	Platform	Classification	Features	No. of subjects	Performance result
Miluzzo et al. [66]	Accelerometer, bluetooth, audio sensors, GPS	Sitting, standing, walking, running, audio info	Social networking	Nokia N95	Decision tree	DFT, FFT and its feature vectors for audio; mean, std. dev. and no. of peaks per unit time deviation of DFT power	22	-
Jigsaw, [60]	Accelerometer, microphone, GPS	Cycling, vehicle, running, stationary, walking, voice, non-voice	JigMe (daily diary), GreenSaw (daily calorie expenditure and carbon footprint)	Nokia N95, iPhone	Decision tree, multivariate Gaussian model, SVM, naive Bayes, GMM	Mean, variance, mean crossing rate, spectrum peak, sub-band energy, sub-band energy ratio, spectral entropy, spectral roll-off, spectral flux, bandwidth, spectral centroid, relative spectral entropy, low energy frame rate, etc.	16	91–95 %
Chu et al. [24]	GPS, accelerometer, microphone, camera	Transportation mode, face recognition, driving, vacuuming, clapping, classical music, showering, etc. (microphone)	Offictfit (reminders to lead healthier lifestyles in office), Cocktail party (record names of other people with whom had a conversation	Android (HTC) phones	HMM, KNN, GMM, decision tree	Transportation mode features, MFCC, speeded-up robust features	Ready datasets were used	16– 76 % reduction in latency and energy
BeWell, [52]	GPS, accelerometer and microphone	Driving, stationary, running, walking, voicing, non-voicing	Automated well-being application	Android phones	naive Bayes	Mean, variance, spectral roll-off (voice)	1–5	85–98 % accuracy
Alqassim et al. [7]	Microphone, accelerometer	Breathing, movement patterns, severe, mild, or no sleep apnea	Sleep apnea monitoring	Windows and Android phones	–	–	–	–

Table 2 (continued)

Reference	Sensors	Activities	Application	Platform	Classification	Features	No. of subjects	Performance result
Matic et al. [64]	Wi-Fi	Interpersonal distance and relative body orientation	Social interaction monitoring	Android phones	naive Bayes, linear classifier, SVM	RSSI (Received Signal Strength Indicator)	8	89 % (highest)
Sociable Sense, [82]	Accelerometer, Bluetooth, microphone	Colocation, stationary, moving, speaker identification	Measuring sociability in office environments	Android, Nokia phones	GMM	–	10	28 % more battery life, 6 % less latency per task
Emotion sense, [83]	GPS, microphone, Bluetooth, accelerometer	Speaker recognition, emotion recognition,	Platform for experimental social psychology research	Nokia phones	GMM	–	12–18	71 % accuracy
IODetector, [100]	Light sensors, Magnetism sensors, cell tower signals	Indoor/outdoor switching mobile applications	Improving the performance of	Android phones	HMM	–	19	Above 88 % (precision, recall)

individuals from the sensing data. In [86], body-worn sensors were used to recognize group activities, such as people walking together, queuing on a line. Accelerometers on mobile phones can also used for distributed activity recognition. For instance, people running together can be identified using the acceleration and Bluetooth (for proximity detection) sensors available on the phones. This open topic should further be investigated for different application domains utilizing different sensors.

- *Open Datasets:* One of the fundamental difficulties in activity recognition research is the challenge of comparing the results of different proposals, since they are usually carried with different number of people, different activities, and different mobile platforms. There is an urgent requirement for the collection of open datasets. Although examples of mobile datasets exist, such as Reality Mining [30] and Nokia-Idiap [45] datasets, they do not utilize the sensors, especially the acceleration sensor, for activity recognition. Small datasets such as the CenceMe [66] also exist, but they do not provide ground truth for performance analysis.

- *Persuasion Methods:* As we mentioned in Section 4, finding efficient ways for persuading users for a behavior change is important, especially in health-care and wellness domains. Finding efficient methods and metaphors [32] to motivate the users is still under investigation.

- *Phone Context Problem:* The location and the orientation of the phone where it is carried is a fundamental challenge in mobile activity recognition. Although the phone context problem has been attempted to be solved by very recent studies [67, 73], there still remain open issues such as the phone context identification in real time including different poses and for different activities (In [73], real-time pose recognition was only tested with the walking activity).

- *Online Classification and Training:* Most of the presented studies benefit from off-line classification methods where the data are collected on the mobile phone but trained and classified off-line on a a backend server. In order to develop real-world applications, classification of activities should be performed online on the mobile platform especially for health and well-being applications. Performance of the classifiers utilized in the presented papers should also be evaluated with online recognition. Moreover, in an ideal system, the training of the classifiers can continuously be improved as long as the user collects data. For instance, using active learning approaches [58], the user can be queried to label the data to improve the recognition rates.

- ***Security and Privacy Threats from Sensors:*** Related to security, in [20, 68], it is shown that the location of screen taps on smart phones can be identified from the readings of motion sensors, such as by the accelerometer and gyroscope. Using this information, an attacker can monitor user's inputs, such as keyboard presses and icon taps. Hence, it is an open field how to address the threats arising from the unrestricted access to motion sensors by the activity recognition systems. Similarly, privacy emerges as a big concern in activity recognition applications due to the potential of collecting personal data, particularly if the data reveal a users location and speech [51]. This raises the requirement for developing suitable privacy-preserving mechanisms [28].

7 Conclusion

In this paper, we provided a review of the state-of-the-art studies in activity recognition on mobile phones, especially targeting health-care and well-being applications. By providing background information on the types of sensors used, targeted activities, performance metrics, and the steps of the activity recognition process, we aimed to present a general snapshot of the problem of activity recognition. We identified the fundamental challenges of activity recognition especially on the mobile platform such as the challenge of continuous sensing and running classifiers on the phones, considering the limited resources available on the phones. Following the challenges identified, we explored the recent studies on the topic and summarized the existing work with a taxonomy. According to this survey and taxonomy, we identified that most of the proposals use off-line training and classification, that simple activities of locomotion are usually targeted, and that they are usually user-dependent systems. Along with these lines, we presented a list of open topics for further research such as fusing information from different sensors for better context recognition, developing user-independent systems to eliminate the training burden, and compiling open datasets to compare the results of different proposals.

References

1. Pulse phone: smart-phone app tracks heart rate. http://www.technologyreview.com/news/425758/smart-phone-app-tracks-heart-rate/. Accessed May 2013.
2. Weka machine learning toolkit. http://www.cs.waikato.ac.nz/ml/index.html. Accessed May 2013.
3. Zeo sleep manager (2012). http://www.myzeo.com/sleep/. Accessed May 2013.
4. Sleepcycle (2013). A bio-alarmclock. http://www.sleepcycle.com/. Accessed May 2013.
5. Mohd Fikri Azli bin Abdullah, Ali Fahmi Perwira Negara, Shosel Sayeed, Deok-Jai Choi, Kalaiarasi Sonai Muthu (2012). Classification algorithms in human activity recognition using smartphones. World Academy of Science, Engineering and Technology, *68*, 422–430.
6. Ahtinen, A., Huuskonen, P., Häkkilä, J. (2010). Let's all get up and walk to the north pole: design and evaluation of a mobile wellness application. In *Proceedings of the 6th Nordic conference on human-computer interaction: extending boundaries, NordiCHI '10* (pp. 3–12). New York: ACM.
7. Alqassim, S., Ganesh, M., Khoja, S., Zaidi, M., Aloul, F., Sagahyroon, A. (2012). Sleep apnea monitoring using mobile phones. In *e-Health networking, applications and services (Healthcom), 2012 IEEE 14th International Conference on* (pp. 443–446).
8. Alvarez, G.G., & Ayas, N.T. (2004). The impact of daily sleep duration on health: a review of the literature. *Progress in Cardiovascular Nursing*, *19*(2), 56–59.
9. Ananthanarayan, S., & Siek, K. (2012). Persuasive wearable technology design for health and wellness. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on* (pp. 236–240).
10. Anderson, I., Maitland, J., Sherwood, S., Barkhuus, L., Chalmers, M., Hall, M., Brown, B., Muller, H. (2007). Shakra: tracking and sharing daily activity levels with unaugmented mobile phones. *Mobile Networks and Applications*, *12*(2–3), 185–199.
11. Arase, Y., Ren, F., Xie, X. (2010). User activity understanding from mobile phone sensors. In *Ubicomp '10: Proceedings of the 12th ACM international conference adjunct papers on ubiquitous computing* (pp. 391–392). New York: ACM.
12. Arteaga, S. (2010). Persuasive mobile exercise companion for teenagers with weight management issues. SIGACCESS Access. *Accessibility and Computing*, *96*, 4–10.
13. Avci, A., Bosch, S., Marin-Perianu, M., Marin-Perianu, R., Havinga, P. (2010). Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: a survey. In *23th International Conference on Architecture of Computing Systems, ARCS 2010* pp. 167–176.
14. Steele, B.G., Belza, B., Cain, K., Warms, C., Coppersmith, J., Howard, J. (2003). Bodies in motion: monitoring daily activity and exercise with motion sensors in people with chronic pulmonary disease. *Rehabilitation Research and Development*, *40*(5), 404–412.
15. Bao, L., & Intille, S.S. (2004). Activity recognition from user-annotated acceleration data. In *Proceedings of Pervasive Conference* pp. 1–17.
16. Berchtold, M., Budde, M., Gordon, D., Schmidtke, H., Beigl, M. (2010). Actiserv: activity recognition service for mobile phones. In *2010 International symposium on wearable computers (ISWC)* pp. 1–8. doi:10.1109/ISWC.2010.5665868.
17. Bieber, G., Koldrack, P., Sablowski, C., Peter, C., Urban, B. (2010). Mobile physical activity recognition of stand-up and sit-down transitions for user behavior analysis. In *Proceedings of the 3rd international conference on pervasive technologies related to assistive environments, PETRA '10* (Vol. 50, p. 5). New York: ACM. doi:10.1145/18392941839354.
18. Bishop, C.M. (2006). *Pattern recognition and machine learning (information science and statistics)*. New York: Springer.
19. Blanke, U., Larlus, D., Laerhoven, K., Schiele, B. (2010). Standing on the shoulders of other researchers—a position statement. In *Pervasive workshop—how to do good activity recognition research? Experimental methodologies, evaluation metrics and reproducibility issues*. Helsinki, Finland.

20. Cai, L., & Chen, H. (2011). TouchLogger: inferring keystrokes on touch screen from smartphone motion. In *Proceedings of the 6th USENIX conference on hot topics in security, HotSec'11* (pp. 9–9). USENIX Association Berkeley.

21. Choudhury, T., Consolvo, S., Harrison, B., Hightower, J., LaMarca, A., LeGrand, L., Rahimi, A., Rea, A., Bordello, G., Hemingway, B., Klasnja, P., Koscher, K., Landay, J., Lester, J., Wyatt, D., Haehnel, D. (2008). The mobile sensing platform: an embedded activity recognition system. *Pervasive Computing, IEEE, 7*(2), 32–41.

22. Choujaa, D., & Dulay, N. (2008). Tracme: temporal activity recognition using mobile phone data. In *EUC '08: Proceedings of the 2008 IEEE/IFIP International conference on embedded and ubiquitous computing*. IEEE Computer Society.

23. Choujaa, D., & Dulay, N. (2010). Activity recognition using mobile phones: achievements, challenges and recommendations. In *How to do good research in activity recognition: experimental methodology, performance evaluation and reproducibility. Workshop in conjunction with Pervasive 2010*.

24. Chu, D., Lane, N.D., Lai, T.T.T., Pang, C., Meng, X., Guo, Q., Li, F., Zhao, F. (2011). Balancing energy, latency and accuracy for mobile sensor data classification. In *Proceedings of the 9th ACM conference on embedded networked sensor systems, SenSys '11* (pp. 54–67). New York: ACM.

25. Consolvo, S., McDonald, D.W., Toscos, T., Chen, M.Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., Landay, J.A. (2008). *Activity sensing in the wild: a field trial of ubifit garden* (pp. 1797–1806). New York: ACM.

26. Dai, J., Bai, X., Yang, Z., Shen, Z., Xuan, D. (2010). Perfalld: a pervasive fall detection system using mobile phones. In *PerCom workshops*, pp. 292–297.

27. Dantzig, S., Geleijnse, G., Halteren, A. (2012). Toward a persuasive mobile application to reduce sedentary behavior. *Personal Ubiquitous Computing*, 1–10. doi:10.1007/s00779-012-0588-0.

28. De Cristofaro, E., Durussel, A., Aad, I. (2011). Reclaiming privacy for smartphone applications. In *Proceedings of the 2011 IEEE international conference on pervasive computing and communications, PERCOM '11* (pp. 84–92). Washington DC: IEEE Computer Society.

29. Dernbach, S., Das, B., Krishnan, N.C., Thomas, B.L., Cook, D.J. (2012). Simple and complex activity recognition through smart phones. In *Intelligent Environments* (pp. 214–221).

30. Eagle, N., & (Sandy) Pentland, A. (2006). Reality mining: sensing complex social systems. *Personal Ubiquitous Computer, 10*(4), 255–268.

31. Fitz-Walter, Z., & Tjondronegoro, D. (2009). Simple classification of walking activities using commodity smart phones. In *OZCHI '09 Proceedings of the 21st annual conference of the australian computer-human interaction special interest group: design: open 24/7* (pp. 409–412).

32. Fogg, B. (1998). Persuasive computers: perspectives and research directions. In *Proceedings of the SIGCHI conference on human factors in computing systems, CHI '98* (pp. 225–232). New York: ACM/Addison-Wesley Publishing Co.

33. Fogg, B. (2007). *Mobile Persuasion: 20 perspectives on the future of behavior change.* Stanford Captology Media.

34. Fontecha, J., Navarro, F., Hervas, R., Bravo, J. (2012). Elderly frailty detection by using accelerometer-enabled smartphones and clinical information records. *Personal and Ubiquitous Computing*. doi:10.1007/s00779-012-0559-5.

35. Garcia Wylie, C., & Coulton, P. (2009). Persuasive mobile health applications. In *Electronic Healthcare.* vol. 0001, (pp. 90–97). Berlin: Springer.

36. Hariharan, R., Krumm, J., Horvitz, E. (2005). Web-enhanced GPS. In *Proceedings of the first international conference on location- and context-awareness, LoCA'05* (pp. 95–104). Heidelberg: Springer.

37. Henpraserttae, A., Thiemjarus, S., Marukatat, S. (2011). Accurate activity recognition using a mobile phone regardless of device orientation and location. In *Proceedings of the 2011 international conference on body sensor networks, BSN '11* (pp. 41–46). Washington DC: IEEE Computer Society.

38. Hynes, M., Wang, H., Mccarrick, E., Kilmartin, L. (2011). Accurate monitoring of human physical activity levels for medical diagnosis and monitoring using off-the-shelf cellular handsets. *Personal Ubiquitous Computer, 15*(7), 667–678.

39. Iso, T., & Yamazaki, K. (2006). Gait analyzer based on a cell phone with a single three-axis accelerometer. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services, MobileHCI '06* (pp. 141–144). New York: ACM.

40. Liu, J., Goraczko, M., Kansal, A., Lymberopoulos, D., Nath, S., Priyantha, B. (2010). Subjective sensing: mission statement and a research agenda. *Technical Report*.

41. Ju, Y., Lee, Y., Yu, J., Min, C., Shin, I., Song, J. (2012). Symphoney: a coordinated sensing flow execution engine for concurrent mobile sensing applications. In *Proceedings of the 10th ACM conference on embedded network sensor systems, SenSys '12* (pp. 211–224). New York: ACM.

42. Kaspar, R., Oswald, F., Wahl, H.W., Voss, E., Wettstein, M. (2012). Daily mood and out-of-home mobility in older adults: does cognitive impairment matter? *Journal of Applied Gerontology*. doi:10.1177/0733464812466290.

43. van Kasteren, T., Alemdar, H., Ersoy, C. (2011). Effective performance metrics for evaluating activity recognition methods. In *ARCS 2011 VDE*.

44. van Kasteren, T., Noulas, A., Englebienne, G., Kröse, B. (2008). Accurate activity recognition in a home setting. In *Proceedings of the 10th international conference on Ubiquitous computing, UbiComp '08* (pp. 1–9). New York: ACM.

45. Kiukkonen, N.J.B., Dousse, O., Gatica-Perez, D. J. (2010). Towards rich mobile phone data sets: Lausanne data collection campaign. In *Proceedings ACM International Conference on Pervasive Services (ICPS2010)*.

46. Könönen, V., Mäntyjärvi, J., Similä, H., Pärkkä, J., Ermes, M. (2010). Automatic feature selection for context recognition in mobile devices. *Pervasive and Mobile Computing, 6*(2), 181–197.

47. Kose, M. (2012). Performance evaluation of classification methods for online activity recognition on smart phones. *Technical Report*.

48. Kose, M., Incel, O.D., Ersoy, C. (2012). Online human activity recognition on smart phones. *Workshop on mobile sensing: from smartphones and wearables to big data (colocated with IPSN)*, 11–15.

49. Krumm, J., & Horvitz, E. (2004). Locadio: inferring motion and location from Wi-Fi signal strengths. In *MobiQuitous* (pp. 4–13).

50. Kwapisz, J.R., Weiss, G.M., Moore, S.A. (2011). Activity recognition using cell phone accelerometers. *SIGKDD Explore Newsletter, 12*(2), 74–82. doi:10.1145/1.9648971964918.

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

52. Lane, N., Mohammod, M., Lin, M., Yang, X., Lu, H., Ali, S., Doryab, A., Berke, E., Choudhury, T., Campbell, A. (2012). Bewell: a smartphone application to monitor, model and promote wellbeing. *IEEE*. doi:10.4108/icst.pervasivehealth.2011.246161.

53. Lane, N.D., Xu, Y., Lu, H., Hu, S., Choudhury, T., Campbell, A.T., Zhao, F. (2011). Enabling large-scale human activity inference on smartphones using community similarity networks (csn). In *Proceedings of the 13th international conference on Ubiquitous computing, UbiComp '11* (pp. 355–364). New York: ACM.

54. Lara, O.D., & Labrador, M.A. (2012). A survey on human activity recognition using wearable sensors. *Technical Report*. http://www.cse.usf.edu/files/69521327941111Survey.pdf.

55. Lee, Y.S., & Cho, S.B. (2011). Activity recognition using hierarchical hidden Markov models on a smartphone with 3d accelerometer. In *Proceedings of the 6th international conference on hybrid artificial intelligent systems - Volume Part I, HAIS'11* (pp. 460–467). Heidelberg: Springer.

56. Li, F., Zhao, C., Ding, G., Gong, J., Liu, C., Zhao, F. (2012). A reliable and accurate indoor localization method using phone inertial sensors. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, UbiComp '12* (pp. 421–430). New York: ACM.

57. Lockhart, J.W., Pulickal, T., Weiss, G.M. (2012). Applications of mobile activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, UbiComp '12* (pp. 1054–1058). New York: ACM.

58. Longstaff, B., Reddy, S., Estrin, D. (2010). Improving activity classification for health applications on mobile devices using active and semi-supervised learning. In *PervasiveHealth*, (pp 1–7).

59. Loseu, V., Ghasemzadeh, H., Ostadabbas, S., Raveendranathan, N., Malan, J., Jafari, R. (2010). Applications of sensing platforms with wearable computers. In *Proceedings of the 3rd international conference on pervasive technologies related to assistive environments, PETRA '10* (pp. 53:1–53:5). New York: ACM.

60. Lu, H., Yang, J., Liu, Z., Lane, N.D., Choudhury, T., Campbell, A.T. (2010). The jigsaw continuous sensing engine for mobile phone applications. In *Proceedings of the 8th ACM conference on embedded networked sensor systems, SenSys '10* (pp. 71–84). New York: ACM.

61. Lukowicz, P., Pentland, A., Ferscha, A. (2012). From context awareness to socially interactive computing. *IEEE Pervasive Computing 11*(1), 32–41.

62. Martin, H., Bernardos, A., Iglesias, J., Casar, J. (2012). Activity logging using lightweight classification techniques in mobile devices. *Personal and Ubiquitous Computing*, *17*, 1–21.

63. Matic, A., Osmani, V., Maxhuni, A., Mayora, O. (2012). Multimodal mobile sensing of social interactions. In *Pervasive computing technologies for healthcare (PervasiveHealth), 2012 6th international conference on* (pp. 105–114).

64. Matic, A., Osmani, V., Mayora-Ibarra, O. (2012). Analysis of social interactions through mobile phones. *Mobile Networks and Applications*, *17*, 808–819.

65. Miluzzo, E., Cornelius, C.T., Ramaswamy, A., Choudhury, T., Liu, Z., Campbell, A.T. (2010). Darwin phones: the evolution of sensing and inference on mobile phones. In *Proceedings of the 8th international conference on Mobile systems, applications, and services, MobiSys '10* (pp. 5–20). New York: ACM.

66. Miluzzo, E., Lane, N.D., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S.B., Zheng, X., Campbell, A.T. (2008). Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application. In *Proceedings of the 6th ACM conference on Embedded network sensor systems, SenSys '08* (pp. 337–350). New York: ACM.

67. Miluzzo, E., Papandrea, M., Lane, N., Lu, H., Campbell, A. (2010). Pocket, bag, hand, etc.-automatically detecting phone context through discovery. In *Proceedings of the first international workshop on sensing for app phones (PhoneSense10)*.

68. Miluzzo, E., Varshavsky, A., Balakrishnan, S., Choudhury, R.R. (2012). Tapprints: your finger taps have fingerprints. In *Proceedings of the 10th international conference on Mobile systems, applications, and services, MobiSys '12* (pp. 323–336). New York: ACM.

69. Miluzzo, E., Wang, T., Campbell, A.T. (2010). Eyephone: activating mobile phones with your eyes. In *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds, MobiHeld '10* (pp. 15–20). New York: ACM.

70. Mladenov, M., & Mock, M. (2009). A step counter service for java-enabled devices using a built-in accelerometer. In *Proceedings of the 1st international workshop on context-aware middleware and services: affiliated with the 4th international conference on communication system software and middleware (COMSWARE 2009), CAMS '09* (pp. 1–5). New York: ACM.

71. Mun, M., Estrin, D., Burke, J., Hansen, M. (2008). Parsimonious mobility classification using GSM and WIFI traces. In *Proceedings of the 5th workshop on embedded networked sensors (HotEmNets)*.

72. Natale, V., Drejak, M., Erbacci, A., Tonetti, L., Fabbri, M., Martoni, M. (2012). Monitoring sleep with a smartphone accelerometer. *Sleep and Biological Rhythms*, *10*(4), 287–292.

73. Park, J.-g., Patel, A., Curtis, D., Ledlie, J., Teller, S. (2012). Online pose classification and walking speed estimation using handheld devices. In *Proceedings 14th international conference on ubiquitous computing (UbiComp 2012)* (pp.113–122).

74. Park, K., Shin, H., Cha, H. (2013). Smartphone-based pedestrian tracking in indoor corridor environments. *Personal and Ubiquitous Computing*, *17*, 359–370.

75. Patrick, K., Griswold, W.G., Raab, F., Intille, S.S. (2008). Health and the mobile phone. *American Journal of Preventive Medicine*, *35*(2), 177–181.

76. Peebles, D., Lu, H., Lane, N.D., Choudhury, T., Campbell, A.T. (2010). Community-guided learning: exploiting mobile sensor users to model human behavior. In *AAAI*.

77. Pentland, A. (2000). Looking at people: sensing for ubiquitous and wearable computing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*, 107–119.

78. Pollak, J., Gay, G., Byrne, S., Wagner, E., Retelny, D., Humphreys, L. (2010). It's time to eat! Using mobile games to promote healthy eating. *Pervasive Computing, IEEE*, *9*(3), 21–27.

79. Preece, S.J., Goulermas, J.Y., Kenney, L.P.J., Howard, D., Meijer, K., Crompton, R. (2009). Activity identification using body-mounted sensors: a review of classification techniques. *Physiological Measurement*, *30*(4). R1.

80. Priyantha, B., Lymberopoulos, D., Liu, J. (2010). Enabling energy efficient continuous sensing on mobile phones with littlerock. In *Proceedings of the 9th ACM/IEEE international conference on information processing in sensor networks, IPSN '10* (pp. 420–421). New York: ACM.

81. Purpura, S., Schwanda, V., Williams, K., Stubler, W., Sengers, P. (2011). Fit4life: the design of a persuasive technology promoting healthy behavior and ideal weight. In *Proceedings of the 2011 annual conference on human factors in computing systems, CHI '11* (pp. 423–432). New York: ACM.

82. Rachuri, K.K., Mascolo, C., Musolesi, M., Rentfrow, P.J. (2011). Sociablesense: exploring the trade-offs of adaptive sampling and computation offloading for social sensing. In *Proceedings of the 17th annual international conference on Mobile computing and networking, MobiCom '11* (pp. 73–84). New York: ACM.

83. Rachuri, K.K., Musolesi, M., Mascolo, C., Rentfrow, P.J., Longworth, C., Aucinas, A. (2010). Emotionsense: a mobile phones based adaptive platform for experimental social

psychology research. In *Proceedings of the 12th ACM international conference on Ubiquitous computing, Ubicomp '10* (pp. 281–290). New York: ACM.

84. Rai, A., Yan, Z., Chakraborty, D., Wijaya, T.K., Aberer, K. (2012). Mining complex activities in the wild via a single smartphone accelerometer. In *Proceedings of the sixth international workshop on knowledge discovery from sensor data, SensorKDD '12* (pp. 43–51). New York: ACM.

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

86. Roggen, D., Wirz, M., Tröster, G., Helbing, D. (2011). Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods. CoRR abs/1109.1664.

87. Ryder, J., Longstaff, B., Reddy, S., Estrin, D. (2009). Ambulation: a tool for monitoring mobility patterns over time using mobile phones. In *Proceedings of the 2009 international conference on computational science and engineering - Volume 04, CSE '09* (pp. 927–931). Washington DC: IEEE Computer Society.

88. Saponas, T.S., Lester, J., Froehlich, J., Fogarty, J., Landay, J. (2008). ilearn on the iphone: Real-time human activity classification on commodity mobile phones. *Technical Report*. UW-CSE-08-04-02.

89. Serra, A., Carboni, D., Marotto, V. (2010). Indoor pedestrian navigation system using a modern smartphone. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, MobileHCI '10* (pp. 397–398). New York: ACM. doi:10.1145/18516001851683.

90. Siek, K. (2012). Persuasive wearable technology design for health and wellness. In *Proceedings of pervasive computing technologies for healthcare (PervasiveHealth), 2012 6th international conference on* (pp. 236–240).

91. Siirtola, P., & Röning, J. (2012). Recognizing human activities user-independently on smartphones based on accelerometer data. *International Journal of Interactive Multimedia and Artificial Intelligence*, *1*(5), 38–45.

92. Sohn, T., Varshavsky, A., Lamarca, A., Chen, M., Choudhury, T., Smith, I., Consolvo, S., Hightower, J., Griswold, W., de Lara, E. (2006). Mobility detection using everyday gsm traces. In *UbiComp 2006: Proceedings of the 8th international conference of ubiquitous computing* (pp. 212–224).

93. Torun, M., van Kasteren, T., Incel, O.D., Ersoy, C. (2012). Complexity versus page hierarchy of a GUI for elderly homecare applications. *Lecture Notes in Computer Science*, *7382*, 689–696.

94. Wang, Y., Lin, J., Annavaram, M., Jacobson, Q.A., Hong, J., Krishnamachari, B., Sadeh, N. (2009). A framework of energy efficient mobile sensing for automatic user state recognition. In *Proceedings of the 7th international conference on Mobile systems, applications, and services, MobiSys '09* (pp. 179–192). New York: ACM.

95. Yan, Z., Chakraborty, D., Misra, A., Jeung, H., Aberer, K. (2012). Semantic activity classification using locomotive signatures from mobile phones. *Technical Report*.

96. Yan, Z., Subbaraju, V., Chakraborty, D., Misra, A., Aberer, K. (2012). Energy-efficient continuous activity recognition on mobile phones: an activity-adaptive approach. In *Proceedings of the 16th international symposium on wearable computers (ISWC)*.

97. Yang, J. (2009). Toward physical activity diary: motion recognition using simple acceleration features with mobile phones. In *Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics IMCE '09* (pp. 1–10).

98. Yavuz, G., Kocak, M., Ergun, G., Alemdar, H.O., Yalcin, H., Incel, O.D., Ersoy, C. (2010). A smartphone based fall detector with online location support. In *Proceedings of PhoneSense 2010*.

99. Young, M. (2010). Twitter me: using micro-blogging to motivate teenagers to exercise. In R. Winter, J. Zhao, S. Aier (Eds.), *Global perspectives on design science research, Lecture Notes in Computer Science* (Vol. 6105, pp. 439–448). New York: Springer.

100. Zhou, P., Zheng, Y., Li, Z., Li, M., Shen, G. (2012). Iodetector: a generic service for indoor outdoor detection. In *Proceedings of the 10th acm conference on embedded network sensor systems, SenSys '12* (pp. 361–362). New York: ACM.