A Survey on Smartphone-Based Systems for Opportunistic User Context Recognition

SEYED AMIR HOSEINI-TABATABAEI, ALEXANDER GLUHAK, and RAHIM TAFAZOLLI,
University of Surrey

27

The ever-growing computation and storage capability of mobile phones have given rise to mobile-centric context recognition systems, which are able to sense and analyze the context of the carrier so as to provide an appropriate level of service. As nonintrusive autonomous sensing and context recognition are desirable characteristics of a personal sensing system; efforts have been made to develop opportunistic sensing techniques on mobile phones. The resulting combination of these approaches has ushered in a new realm of applications, namely *opportunistic user context recognition with mobile phones*.

This article surveys the existing research and approaches towards realization of such systems. In doing so, the typical architecture of a mobile-centric user context recognition system as a sequential process of *sensing*, *preprocessing*, and *context recognition* phases is introduced. The main techniques used for the realization of the respective processes during these phases are described, and their strengths and limitations are highlighted. In addition, lessons learned from previous approaches are presented as motivation for future research. Finally, several open challenges are discussed as possible ways to extend the capabilities of current systems and improve their real-world experience.

Categories and Subject Descriptors: A.1 [**General**]: Introductory and Survey; I.5.2 [**Pattern Recognition**]: Design Methodology—*Classifier design and evaluation, Feature evaluation and selection*; I.2.6 [**Artificial Intelligence**]: Learning—*Knowledge acquisition*

General Terms: Algorithms, Design, Experimentation, Performance

Additional Key Words and Phrases: Pervasive computing, opportunistic sensing, smartphone, user context recognition

ACM Reference Format:
Hoseini-Tabatabaei, S. A., Gluhak, A., and Tafazolli, R. 2013. A survey on smartphone-based systems for opportunistic user context recognition. ACM Comput. Surv. 45, 3, Article 27 (June 2013), 51 pages.
DOI: http://dx.doi.org/10.1145/2480741.2480744

1. INTRODUCTION

Efforts to understand human behavior are well-documented since the early physiologica, psychological, and sociologic studies of the 18th and 19th centuries. Since then, different branches of science with different perspectives have studied human behavior in terms of relations between different causes, events, and types of behavior. A brief look at scientific theories about humans shows that the causes of the behavior involve biological aspects such as hormonal state or genetic inheritance; sociological aspects such as social esteem, gender, culture, and religion; mental aspects such as IQ or cognition; and many other causes and scientific factors [Martin et al., 2007]. Given that human behavior is rooted in the combination of these causes, a single perspective can never give a comprehensive explanation of behavior. When we add to this fact

Corresponding author's email: S.Hoseinitabatabaei@surrey.ac.uk.

DOI: http://dx.doi.org/10.1145/2480741.2480744

the uniqueness of an individual, understanding human behavior from its internal and personal cause and effect perspective appears to be an unattainable goal. A solution to this problem may be to focus on the external effects of these causes in an individual's daily life by developing an understanding of their behavior based upon the correlation between what individuals expose to their environment and a specific type of behavior. Observation of such correlations can thus be utilized to develop a model for human behavior in various situations. Studies in human behavior show that a person's behavior is highly dependent on perception, context, environment, prior knowledge, and interaction with others [Atallah and Yang 2009]. In this regard, various studies (e.g., [Atallah and Yang 2009; Aoki et al. 2002; Suh et al. 2009]) have concluded that in order to model human behavior, a complete context of the human's activities, interactions, and surrounding environment is required. These contexts are sometimes referred to as spatial, personal, and social aspects [Suh et al. 2009] or *user context* (UC) in context aware systems [Mostefaoui et al. 2004].

Recent advances in the semiconductor industry and in wireless communications have contributed to the development of alternative observation capabilities based on a variety of miniaturized sensors and computing technologies. These are gradually replacing the old-fashioned questionnaires, surveys, and participatory observation techniques traditionally used to capture such information. Ambient sensors and body sensor networks (BSN) have been typically used for sensing different aspects of a user's context. However, these technologies are typically suited for observations in limited geographic scope and over short periods of time, due to the dependency of the ambient sensors on the infrastructure and the intrusiveness of the BSNs. Real-world applications of ambient and wearable sensor observations were consequently limited to surveillance, analyzing behavior of a group of participants during a study, or healthcare monitoring, where patients would accept wearing the device for a long period of time. The collection of longterm user context information with ubiquitous coverage still remains an open technological challenge.

In light of new advances in computing, storage, and wireless technology and the recent introduction of MEMS (micro electro mechanical system) sensors into mobile phones, a door to a new world of application possibilities has been opened. Given the indispensible role of the mobile-phones in everyday life, mobilephone-centric sensing systems are ideal candidates for ubiquitous observation. The current applications of pervasive mobile phone sensing primarily include the reproduction of the healthcare approaches using BSNs, modeling user movement patterns, environmental monitoring, and discovering social interactions. With respect to human-centric sensing, mobile-phone-based sensing and wireless sensor networks, in particular BSN-based approaches, share many similar research challenges, such as energy, security, and privacy. Techniques developed for either of these systems are applicable to both. The combination of BSNs and mobile technology has attracted many researchers to develop applications in mobile phones that process the data gathered from a BSN. To differentiate between previous work in wireless sensors and particularly in the field BSNs, the primary focus of this study is on the methodologies where the entire process from sensing to recognizing the various aspects of user context is performed on a smartphone. In such methodologies, the mobile embedded sensors are used for data acquisition, while the computational capability of mobile phones is exploited for user context recognition through a sequential data processing architecture. After an initial sensing or data acquisition phase, the sequence of processes typically consists of a preprocessing and a context inference phase. We introduce these phases and their interactions in Section 1.1 and then extensively investigate their related techniques and issues in the remaining part of this work.

The selection of required computational techniques strongly depends upon the level of active user involvement in the sensing process [Lane et al. 2010]. Approaches that are supported by the active involvement of the user, for example, by providing explicit input or decisions to the sensing process, are called *participatory sensing*. In contrast, methods that operate autonomously without user involvement are more challenging and are referred to as *opportunistic sensing*. More details on these aspects are provided in Section 1.2.

This work provides a survey of the state-of-the-art techniques for opportunistic mobile-centric user context recognition systems. There are three objectives of this work. The first is to classify the current methodologies in opportunistic phone sensing as different components of a mobile sensing architecture. To our knowledge, this literature is the first survey that has provided such information about this domain. The second is to provide an overview and analysis of the more recent progress made toward solving the key challenges for realizing opportunistic sensing systems. This will allow researchers to better understand the currently available capabilities. The third objective is to present several remaining issues and possible future directions of this research area.

The remainder of this article is organized according to the architecture of mobile-phone-centric user context recognition systems. Section 2 will discuss the sensors embedded in current mobile phones and their respective sensing capabilities. Section 3 focuses on preprocessing, discussing recent advances and techniques for calibration, and feature extraction. Section 4 investigates the currently widely implemented algorithms for the context recognition phase and analyzes their computational characteristics. Section 5 summarizes all of the mentioned aspects and provides a comprehensive overview of the latest applications. Finally, Section 6 highlights some of the future challenges and opportunities in related fields.

It should be mentioned that, apart from using data from mobile embedded sensors, researchers have explored a variety of different data sources from mobile phones for modeling and understanding different facets of human behavior. Examples range from analysis of the pattern of message communication and phone calls (e.g., [Fawcett and Provost 1996; Vieira et al 2010]) to logs of Internet browsing data (e.g., [Olmedilla et al. 2010]) and application usage for calendar, music, or photo browsing (e.g., [Zulkefly and Baharudin 2009]). However, the respective analyses are usually performed offline and in backend servers, and so do not fit the scope of this article.

1.1. Mobile-Phone-Centric User Context Recognition

The potential of exploiting mobile phones for sensing and context recognition research has long attracted researchers in both industrial [Nokia 2005] and academic research communities [Eagle and Pentland 2006]. However, the majority of advances has taken place only recently. In their recent survey on mobile phone sensing, Lane et al. [2010] argue that the recent acceleration of progress in this field is the result of four main technological advances.

(1) The presence of low-cost and powerful sensors in mobile phone devices.
(2) The facilitation of the entrance of third-party programmers by offering them software development kits (SDK) and application programming interfaces (APIs).
(3) The introduction of application stores that enables developers to deliver their applications to a large number of users across the world.
(4) The mobile computing cloud that Enables developers to take advantage of resources in back end servers as well as for analyzing and collecting data from a large number of users.

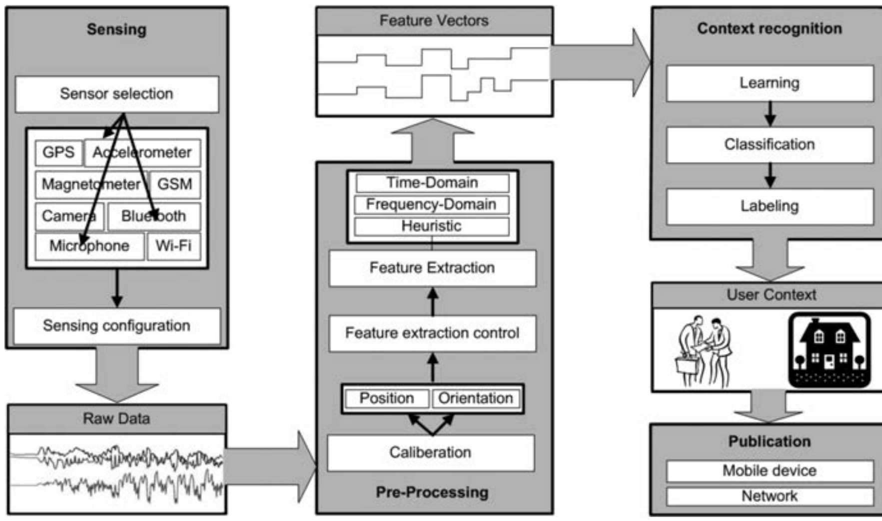

Fig. 1. Overview of tasks and data flow of mobile phonecentric sensing for user context recognition.

The combination of these factors has accelerated the rise of innovative mobile sensing applications, which are likely to lead to a revolution in everyday life in the near future. Early examples of such successful and popular applications are SenSay [Siewiorek et al. 2003], Micro-Blog [Gaonkar et al. 2008], PeopleNet [Motani et al. 2005], MyExperience [Froehlich et al. 2007], Serendipity [Eagle and Pentland 2005], Place-its [Sohn et al. 2005] and CenceMe [Miluzzo et al. 2008].

Systems for user context inference on mobile phones rely on a variety of technologies from different domains, including artificial intelligence, digital signal possessing, human computer interactions, and ubiquitous computing. Since sensing with mobile phones is still in its infancy, no clear consensus on sensing architecture on mobile phones currently exists [Lane et al. 2010]. Our survey provides an important step in this direction, by reviewing the recent advances in mobile-based sensing and identifying the essential aspects that have been recently proposed in the different existing approaches. Mobile-phone-based user context recognition methodologies typically realize a sequence of main system stages, as shown in Figure 1.

The initial sensing step typically produces raw observational and measurement data that is often refined in a preprocessing step. The refined data or features extracted through preprocessing are then passed to context inference processes before the measured context is delivered to the context consumer (i.e., an application on the mobile phone or backend server). The energy constraints on the mainly batterypowered portable handsets make the configuration of sensing very challenging. The main challenge that impacts the sensing stage is to accurately recognize the required context with a minimum number of sensors and sensing frequency. Through preprocessing, the phone's context ambiguity is resolved via a calibration process prior further processing steps. The constraints of computation and memory resources also limit the implementations of preprocessing and classification techniques to less computationalintensive methods. During the preprocessing stage, redundancy and noise are minimized in the raw data and more computationally efficient representations of the data (called features) are derived. Features are used as inputs to the classification techniques that determine the computed user context. The selection of features for a classifier is often performed offline, by frequent training and evaluation of classification performance, with the aim of improving the classification accuracy by discarding indiscriminative or highly correlated features and avoiding the curse of dimensionality. The availability of certain features depends on a phone's context, which is often dynamic. Furthermore, the extraction of features may require techniques of different computational complexity. Therefore existing mobile sensing system architectures require feature extraction control mechanisms to exploit these trade-offs dynamically, in order to better adapt to the resource constraints of the mobile sensing platform and varying context requirements.

Finally, the derived context (or sensed data) is delivered to either a backend server or to an application on the mobile phone for consumption. Delivering the context to locally-consuming services and applications on a mobile phone causes fewer privacy concerns and reduces the power required for transferring the data. However, the complexity of the applications is limited by the available local computing and storage resources. Uploading data to a backend server provides better opportunities for the exploitation of aggregate data from a large number of users, while allowing for the realization of more complex applications. However, it demands more serious considerations of privacy, and requires higher power consumption for the remote context delivery.

1.2. User Involvement in the Sensing Process

Based upon the level of user involvement during the sensing process, the sensing applications can be divided into *participatory sensing* applications, where the user is actively participating in the sensing process, or *opportunistic sensing,* where the user remains passive and is not required to participate. While the identified system components introduced in the previous section essentially apply for both of these categories, the techniques for realizing the system stages, from sensing to context inference, can differ.

In participatory sensing, complex operations can be supported by leveraging the intelligence of the user, which can significantly reduce the sensing, calibration, and classification challenges when compared to an opportunistic approach. For example, the information about the orientation and the position of the device or identifying the user's context can be directly provided (or at least corrected) by the user. This can significantly reduce the computational requirements on the device. More importantly, with user supervision, there will be an increased user awareness regarding the contents of the sensed data. This eventually improves the acceptability of this approach in terms of privacy. Despite several advantages of participatory approaches, some drawbacks must also be considered. One particular drawback is that the data specifications and characteristics, for example, time, duration, location, space, contents, etc., are dependent upon participants' enthusiasm and willingness to collect data during their daily lives. Moreover, collected data are affected by a bias from the user's knowledge or opinions during the data collection. The problematic effects of this fact are well known and carefully considered in data collection methods for human subject studies [McNeill and Chapman 2005].

Opportunistic sensing, alternatively, lowers the burden placed on the user which in return implies that collected data is less affected by user characteristics. One of the key challenges in opportunistic sensing systems is determining how to transfer the required sensing functionality and intelligence to mobile phones without jeopardizing the phone experience with the overhead of additional processing. For example, the position of the phone relative to the user's body is a key parameter for activity recognition. With the lack of user participation, such algorithms require the execution of a calibration process that automatically identifies the device position prior to activity recognition, adding significant computational burden to the mobile phone, thus shortening battery life. The classification methods, while being computationally simple, must be able to accurately recognize the user context and even cope with the presence of unknown contexts, thereby providing scalability in the methods' context recognition techniques.

Generally speaking, these systems are often technically more difficult to realize [Das et al. 2010] but provide more reliable data and tend to attain more acceptance by users, since their application is less intrusive.

2. SENSING

Sensors available on mobile phones can be classified as inertial, positioning, and ambient sensors. Each of these types of sensors is capable of sensing different aspects of user context and are selected and configured based upon application requirements. In this section, these sensors, their sensing capability, and current applications are introduced. In addition, Table II in the Appendix shows how they have been utilized in different mobilephonecentric context recognition systems.

2.1. Inertial Sensors

Inertial sensors are usually referred to as sensors that are able to measure the physical motion of a solid object. Recently, mobile phones have been equipped with inertial sensors, such as accelerometers and gyroscopes. Their characteristics and applications are described in the following sections.

2.1.1. Accelerometers. Accelerometers are typically electromechanical instruments that measure the applied acceleration acting along their sensitive axis. The measured acceleration can be static, like the constant force of gravity, or dynamic, caused by moving or shaking the accelerometer. Regardless of manufacturing and design differences, the accelerometer's functionality is a variation of a spring/mass system. In this system, the acceleration is proportional to the displacement of the mass when the force is applied. MEMS-based accelerometers have long been used as a primary resource for capturing context information with wearable technologies [Yi et al. 2005]. Examples of such research are relative positioning systems (a.k.a. dead reckoning) (e.g., [Levi and Judd 1996; Olguin and Pentland 2006]), pervasive activity recognition applications such as physical work monitoring [Stiefmeier et al. 2008], health care applications such as estimating energy expenditure, fall detection, and activity level (e.g., [Redmond and Hegge 1985; Bouten et al. 1997; Wu et al. 2007; Choudhury et al. 2008; Lester et al. 2006], and ambulatory monitoring (for extensive discussion in this field, see [Mathie et al. 2004]). Developing such applications requires the ability to discriminate between different user physical activities contained within the accelerometer data ranging from coarser levels, such as moving or stationary modes for dead reckoning approaches, to finer levels of movement, such as running, walking, sitting, or standing, and even the transition patterns between them in healthcare approaches. It has been successfully verified in many studies (e.g., [Ravi et al. 2005; Bouten et al. 1997; Choudhury et al. 2008]) that a single accelerometer attached to the user body is enough to detect the majority of daily life activities with the accuracy required for these applications. Accelerometers are also found in many smartphones. Their primary purpose is to detect changes in the orientation of the mobile phone, so as to rotate the screen's display or captured images in accordance with the phone's orientation. Recent studies have used these accelerometers for detecting the user's physical activities while carrying a mobile phone. However, it is unclear as to what extent these embedded accelerometers are capable of detecting a user's activity. The following discussion will explore this issue further.

The acceleration generated during human movement varies across the body and depends upon the activity being performed. This acceleration increases in magnitude from the head to the ankle and is generally greatest in the vertical direction [Bhattacharya et al. 1980]. Despite vertical acceleration being the most dominant component, it is inadvizable to neglect the horizontal acceleration [Lafortune 1991].

Mathie et al. [2004] provide a comprehensive analysis of acceleration measurements with respect to different daily activities. According to this study, running produces the greatest vertically-directed acceleration amongst other ordinary daily activities, followed by walking down stairs and jumping on a trampoline, while walking up stairs, walking on level terrain, and cycling produce lower magnitudes of acceleration. For many researchers, the detection of walking activity as the most frequent daily activity [Kunze et al. 2005] is of great importance.

Studies of the range of magnitude and the frequency range of acceleration generated by the body during daily life activities (e.g., [Antonsson and Mann 1985; Aminian et al. 1995; Bouten et al. 1997]) confirm that the capabilities provided by accelerometers embedded in current mobile phones are sufficient for detection of almost the same range of activities as with the current wearable approaches. For example, Cappozzo [1989] stated that during walking, upper body accelerations in the vertical direction have been found to vary from -0.3 to 0.8 g, while Sun and Hill [1993] have found that the major output of work for daily activities is confined to a frequency range from 0.3 to 3.5 Hz. Accelerometers such as the LIS302DL digital output accelerometer [STMicroelectronics 2010] that are embedded in Nokia and Apple smartphones [Yang 2009; Hailes et al. 2009] provide tri-axial measurement with a configurable range of ± 2 g or ± 8 g and an output rate of 100 Hz or 400 Hz [STMicroelectronics 2010]. However, the studies from the wearable community also suggest that the accuracy of the results is strongly dependent upon the position of the accelerometer on the user's body. A further issue is that the claimed sampling frequency of embedded accelerometers is often not achievable in practice on mobile phones due to implementation constraints and restricted access to the full resolution and sampling frequencies by the corresponding APIs [Brezmes et al. 2009]. For example, our experiments with Android-based G1 phones from HTC have demonstrated a realistically achievable sampling frequency range of 5–25 Hz. Similarly, Yang [2009] has reported a maximum usable sampling frequency of only 36 Hz on a Nokia N95 device. The power consumption of accelerometers is very small compared to other sensing modalities, for instance, the LIS302DL power consumption is less than 1 mW.

2.1.2. Gyroscopes. Recently smartphones have been equipped with MEMS Gyro sensors. MEMS gyroscopes are non-rotating sensors which basically use the Coriolis Effect on a mass to detect inertial angular rotation [Titterton and Weston 2002]. The embedded gyro sensors have been used in physical activity recognition (e.g., [Morris and Pradiso 2002]), body posture detection (e.g., [Cho et al. 2004]) and dead-reckoning applications (e.g., [Kourogi and Kuratta 2003]). However, probably the most successful application of the embedded gyroscopes has been in digital camera stabilizing techniques (e.g., [Yong-Xiang et al. 2009] that seem to also to be the primary task of gyros in new mobile phones. MEMS-based gyroscopes are considered to have very low power consumption. However, using the gyro sensors for orientation estimation is prone to error accumulation as a result of significant calibration errors, electronic noise, and temperature [Woodman 2007].

2.2. Positioning Sensor and User Proximity Detection

Contemporary mobile phones comprise a number of sensors capable of sensing the user's location and presence of entities in her proximity. Apart from GPS, which is primarily used for outdoor positioning, GSM, WiFi, and Bluetooth signals are also used for user localization (for extensive readings about ubiquitous localization refer to [Hightower and Borriello 2001]). The short-range communication link that can be provided by Bluetooth devices is also a very popular tool for probing a user's surroundings. This technique has gained the attention of many researchers, including social scientists. In this section, an overview of these technologies is provided, along with some examples of their applicability for mobile-centric sensing.

2.2.1. Bluetooth. Bluetooth is a universal, low-cost interface for ad-hoc wireless connectivity, originally developed by Ericsson in 1994, released in 1998 to operate in 2.4–2.48 GHz band, and later ratified as IEEE standard 802.15.1. Bluetooth is designed for short-range communication. Bluetooth devices of power class 2, which are typically implemented in handsets, are limited to a range of approximately ten meters with a transmission power around 2.45 mW(4 dBm), and version 2.0 (2004) of Bluetooth communication, for example, is capable of transmitting up to 3 Mbit/s [Schiller 2003]. The main application of Bluetooth for sensing purposes has been in logging local devices and communicating with external sensors or services. Bluetooth standard allows devices to performe device discovery so as to obtain information about other Bluetooth-enabled devices in their vicinity. This information includes the Bluetooth MAC address, which is also referred to as a Bluetooth identifier (BTID), the device type, and device name. The BTID is a 48-bit number which is unique to a particular Bluetooth device. A device name is defined by users, and the device type is a set of three integers representing the type of discovered device (laptop, phone, etc.). The ability of Bluetooth to sense the presence of other devices in close proximity to the user has been widely employed in social intelligence applications. The high power consumption of continuous Bluetooth scanning for detecting the proximate devices makes battery life in mobile devices a concern [Crk et al. 2009].

2.2.2. Cell Tower Signals. In a mobile communication network, the geographical region of the network is divided into cells. Each cell is a geographic area within which mobile devices can communicate with a particular base station. A base station is interconnected with other base stations mostly through a wired backbone network, while it communicates with mobile devices in its territory via wireless channels. Mobile phones are continuously receiving signals from nearby cell towers. Depending on a variety of parameters, such as network traffic and signal strength, a phone in a cellular network can be connected to different cell towers in different locations at different times. Logging the proximate tower's ID over time has been widely used as a technique for localizing mobile users (e.g., [Kim and Lee 1996]). Cell tower IDs are uniquely identified by a combination of mobile country code (MCC), mobile network code (MNC), location area code (LAC), and cell identifier [Sohn et al. 2006]. Researchers have also tried to analyze the data from mobile phone operators (e.g., [Gonzalez et al. 2008; Onella et al. 2007]), such as call data records (CDR). However, CDRs provide an estimation of the location only during the time that the device is in use. Therefore, as suggested in Eagle et al. [2009], the only option up until now for obtaining continuous cellular tower data has been to prepare a logging application on the mobile device itself. A mobile device may sense a number of cell towers belonging to the same region but from different network providers. Redundancy in collected data can be minimized by constraining the logging software to only log the cell towers belonging to a particular operator, as in Sohn et al. [2006], or to a particular LAC (e.g., [Anderson and Muller 2006b]). Maintaining mobile-to-base station communication when a user is moving requires the network to provide migration service provision from one cell to another. This process is called a *hand-off* or *hand-over* and typically occurs when the received signals on a mobile phone drop below a predetermined threshold. Varying speeds of user movement poses different distributions of received cell IDs according to the hand-off strategies and the distribution of cells in the user environment, for example, fluctuation of cell IDs in a metropolitan area may have different patterns to ID fluctuations in an urban area. The cell IDs' fluctuation pattern in the company of signal strength fluctuation patterns is widely used for obtaining coarse information about the user's physical activities [Anderson et al. 2007].

2.2.3. GPS. Global Positioning System provides the position of the user nearly anywhere on Earth. GPS is based on simultaneous broadcast signal measurements and comparisons that can be carried out from a mobile unit [Kyriazakos and Karetsos 2000]. The position of a mobile phone can be measured based upon the time of signal delay from each of a number of satellites [Mishra 2004] to a mobile satellite, giving position in two dimensions (latitude, longitude) when the receiver is able to see at least three satellites. Zhao outlines in his study [2000] that civilian applications can exploit GPS signals transmitted at 1575.42 MHz using code-division multiple-access (CDMA) techniques with direct-sequence spread-spectrum (DS-SS) signals at 1.023 MHz (Mchips/s) [Zhao 2000]. A satellite's DS-SS signals include accurate time and coefficients (ephemeris) that describe the satellite's position as a function of time. The ground GPS receiver position is determined by the time of arrival (TOA) of these signals. The horizontal accuracy of this system is between 50 to 80 meters and, by means of differential GPS, can be improved to an accuracy of up to 10 meters [Kyriazakos and Karetsos 2000]. Positioning of mobile users with GPS or GSM signals (which will be introduced later) is especially desirable for network operators, as it allows them to provide a variety of value-added services based upon user location. Kyriazakos and Karetsos [2000] have classified the application of mobile user positioning for operators into a number of services, such as safety, billings, information, tracking, and multimedia. An example of such services can be the NAVITIME application [Arikawa et al. 2007] which helps pedestrians find the best route to their destination based on different parameters, such as weather at the destination and the amount of carbon dioxide the user's transport may emit during the trip. Many researchers have emphasized the opportunity that the use of mobile phone GPS sensors can provide for studying the traveling behavior of users [Yim 2003; Yim and Cayford 2001; Ohmori et al. 2005]. It has even been suggested that mobile GPS data replace conventional survey data gathered about a user's traveling behavior [Ohmori et al. 2005]. Traveling information from mobile devices is used in a variety of applications, such as traffic estimation [Herrera et al. 2010] or helping riders with navigation and driving tips [Barbeau et al. 2010].

Despite the high accuracy of GPS technology for outdoor localization, GPS is usually considered the most power-hungry localization technique for mobile computing [Gaonkar et al. 2008].

2.2.4. WiFi. IEEE 802.11 (WiFi) is a means to provide wireless connectivity to devices that require quick installation or, in general, to mobile devices inside a wireless local area network (WLAN) [Ferro and Potorti 2005]. The spectrum ranges from 2.4 to 2.4835 GHz in the U.S. and Europe, while in Japan it ranges from 2.471 to 2.497 GHz. As compared to Bluetooth—the other widely-available, short-range wireless communication method—WiFi typically provides communication ranges of up to 100 meters but with much higher power consumption (30–100 mW). WiFi connections can also provide higher rates (up to a few hundred Mb/s for one-way data) and they have fewer limitations in terms of the maximum number of devices in a basic cell [Ferro and Potorti 2005]. A comprehensive comparison between Bluetooth and WiFi communication and protocols is provided in Ferro and Potorti [2005]. A WiFi device scans the available channels by sending probe requests in order to discover an active network that, in return, sends probe responses. At this stage, the logging of the MAC address of access points or the SSID (service set identifier) of the network with a known location can be used for localizing the scanning device (e.g., [Bahl and Padmanadhan 2000; Grisworld et al. 2002]). However, due to the larger WiFi signal transmission range, the positioning accuracy is coarser, and so supplementary information, such as signal strength (e.g., [Krumm and Horvitz 2004]) or signal triangulation and fingerprinting when using multiple access points (e.g., [Kansal and Zhao 2007]) or a combination of these (e.g., [Cheng et al. 2005]) have been used. A comparison between GPS, WiFi, AGPS, and GSM localization in Gaonkar et al. [2008] has shown that after GPS, localization techniques based upon the detection of WiFi access points is the most power-demanding approach. As a result, WiFi is typically used as a secondary and complementary instrument while in the company of Bluetooth [Miluzzo et al. 2008] or GSM signals (e.g., [Gaonkar et al., 2008]), for indoor localization techniques.

2.3. Ambient Sensors

As discussed in the previous sections, location sensors and inertial sensors on a mobile device can provide information about the persons who act as their carriers. In this section, we discuss sensors that can be used for sensing the surroundings of a user, such as a camera, magnetometer, or microphone. Exploiting their capabilities for sensing environmental state, some researchers have used a network of mobile phones as a sensor network for environmental monitoring purposes [Kanjo et al. 2009].

2.3.1. Camera. The mobile phone's camera is a ubiquitous imaging device with powerful image capture and processing capabilities. It is therefore not surprising that, in addition to its main function as an image capturing tool, it is also a useful enabler of a variety of additional applications. Examples of these applications include the recognition of objects in museums [Ruf and Detyniecki 2009], [Bruns et al. 2007], gesture recognition (e.g., [Wang et al. 2006], [Haro et al. 2005]) location identification (e.g., [Davis et al. 2006], [Ravi et al. 2005], [Lim et al. 2007]), and document recognition, that is scanning (e.g., [Liu et al. 2006], [Erol et al. 2008]. Usually, these applications require a client/server architecture, where computationally-intensive image processing and classification are carried out on backend servers (e.g., [Lim et al. 2007; Chen et al. 2009]).One example is a study by Takacs et al. [2008] that provides augmented reality on mobile phones, where the camera phone images are processed on the phone to be matched against a large database of location-tagged images on backend servers. Sometimes, picture frames are used directly with no further processing (e.g., [Miluzzo et al. 2008; Larsen and Luniewski 2009]), or simple and computationally-affordable techniques are used directly on the mobile phone [Wagner et al. 2010]. (For a comprehensive discussion, see [Gu et al. 2008].)

Opportunistic sensing with a camera is not as straightforward as it is with these sensors. For instance, since the pictures are not taken under user direction, the data acquisition technique must be able to ensure, with reasonable confidence, that the taken picture contains the proper data about the user's surroundings (e.g., the phone is not in the user's pocket). Moreover, as continuous sampling of a camera can generate larger quantities of data and exhaust available battery reserves, adequate data management techniques are essential for resource-constrained mobile sensing systems.

2.3.2. Magnetometer. Digital compasses are another class of sensors that have gained popularity in mobile phones. At the heart of these solutions are tri-axial vector magnetometer sensors, which are able to sense the magnitude of the surrounding magnetic field along their sensitive axes. The magnetometers embedded in mobile phones typically utilize the Hall effect. The sensed magnetic field is a combination of the effects of the Earth's magnetic field and that of surrounding objects. The effect of surrounding objects can be divided into deterministic interference, which includes the effect of ferrous materials (soft iron) and magnetized materials (hard iron) and nondeterministic interference. The effect of nearby objects can distort or even drown out the weaker direction of the Earth's magnetic field for navigation proposes. In this case, additional systems are required alongside magnetometers for compensation for surrounding interference.

Currently, magnetometers embedded in mobile phones are typically equipped with a Dynamic Offset Estimation (DOE) system for compensating for the deterministic interference. The nondeterministic interference can also be effectively mitigated by proper shielding of the sensor and performing simple filtering over the measured magnetic field [Fang et al. 2005].

Portable sensing of the ambient magnetic field provides opportunities for a variety of applications. Lee and Mase [2002] have used the digital compass for dead reckoning. Statistical analysis of accelerometer, magnetometer, thermometer, and light sensors has also been proposed in Golding and Lesh [1999] for portable indoor navigation systems. In such systems, the direction of movement is detected with the magnetometer, while the accelerometer and gyroscopic sensors are used for gait recognition. A similar approach has been implemented by Lee and Mase [2001, 2002]. Some personal navigation systems for mobile devices combine compass and GPS information. Assuming that the user is aware of the mobile phone's orientation, the system simply provides a comparison between the phone's orientation (with the sensor output from the magnetometer) and the static directions (e.g., North, South, etc.) of a map. Other examples of smartphone-based applications exploiting embedded magnetometers include a three-degrees-of-freedom controller for 3D object rotation tasks based upon innovative techniques, such as those proposed in Katzakis and Hori [2009]. Magnetometers embedded in mobile phones are very efficient in power consumption. For example, the AK8976A device which is used in the HTC Dream handset consumes 6.7 mA during sensor operation and just 460 μA of average current with measurements at 100 ms intervals [Asahi Kasei 2007].

2.3.3. Microphone. A microphone is an acoustic transducer, typically with a conversion of about 10 mV/Pa and a signal-to-noise ratio of about 68 dB for the frequency range of 20 Hz to 10 kHz. In addition to their use in voice calls, researchers have recently tried to develop different applications based upon the sensing capabilities of a mobile phone's microphone. A very successful example is that of speech recognition systems [Deligne et al. 2002], which are widely implemented in current mobile phones. These systems enable users to operate the mobile phone by means of voice command without a keyboard. Pervasive applications based on microphones, as discussed by Choudhury et al. [2008], typically involve recording people in unconstrained and unpredictable situations, both in public and in private. These recordings may involve information that the user may not have intended to share. Therefore, most sensing applications focus on extracting nonverbal features from the recorded sound before sharing any information. Nonverbal vocal cues, such as the pattern of silent moments, pitch, or tempo of the speech, have been used in sociometer badges, such as Meeting Mediator (MM) [Kim et al. 2008], to give feedback about the user's social behavior on his mobile phone. Another example of nonverbal features is the analysis of ambient noise to measure noise pollution in environmental monitoring applications [Kanjo 2010] or for detecting the presence of conversation in context-aware applications.

3. PREPROCESSING

In order to reduce data redundancy, noise, and jitter in instantaneous sensor readings, measured values are usually passed to a preprocessing stage. Figure 2 provides a flow chart of a typical preprocessing stage of mobile phone-centric sensing systems. The preprocessing first filters the raw sensor data by minimizing the errors related to noise or jitter during sensing procedures and calibration problems and then converts it into a set of finite features or categories. Based upon the applied sensors and the required quality of data, many different noise and jitter algorithms have been developed to obtain a consistent data stream. The body of existing work on these algorithms is

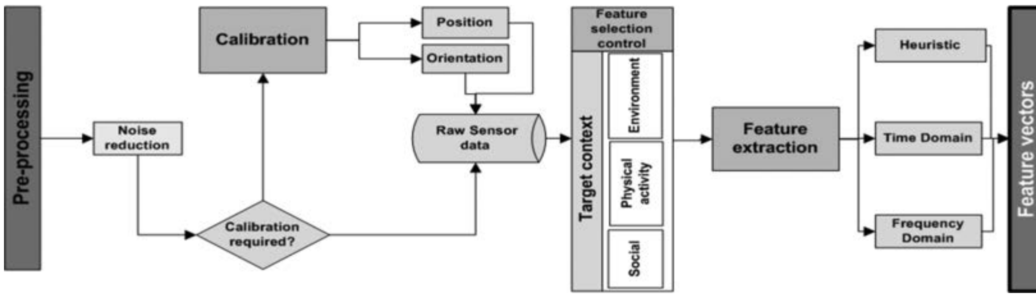

Fig. 2. Overview of different procedures during the preprocessing stage.

immense, even when limited to the aforementioned sensors. The discussion in the first part of this section will be more narrowly focused on the methods of addressing the limitations and errors inherited from the handset properties of a mobile-centric sensing system, namely the calibration, or *phone context* issues [Lane et al. 2010]. Such sensing systems must be prone to frequent changes in orientation and position of the device during data collection, preparation, and feature extraction, while still being able to generate informative and computationally-efficient features. The second part of this section is dedicated to an introduction to diverse feature extraction techniques available for different aspects of a user's context. Despite the discussed problems, which affect all sensing systems that are developed for mobile phones, errors related to the specification of implemented mobile platforms, such as added errors in rough quantization [Bieber et al. 2009], inconsistency in sensor readings [Bieber et al. 2009] and operating system limitations [Miluzzo et al. 2008], are not discussed in this study.

3.1. Calibration

Analogous to Martens and Naes [2002], calibration is defined as a process that enables one to predict an unknown quantity Y from an available observation X through some mathematical transfer function, where the Y value would be the calibrated value at a known reference. Compared to other sensing systems, which consider a fixed position and orientation for their sensors, mobile phones are carried and used in ways that are difficult to anticipate in advance for a particular user. The outputs of inertia sensors and ambient sensors are susceptible to the phone position and/or orientation. For instance, the quality of sound and picture samples is susceptible to the position of the sensing device (e.g., the phone could be in the user's pocket or hand). Thus, adding the orientation information to samples from the camera could help to provide more informative features.[1] Therefore, providing a pervasive sensing system on a mobile phone requires a calibration process in order to transfer the measured data into known location and orientation references. These references are predefined positions and orientations of the device that are used in feature extraction and subsequently the learning process of classifications methods (described in Section 4). An analysis of literature on opportunistic sensing based on microphones and cameras shows that simple heuristic techniques are typically adequate for addressing the required information about the phone position (e.g., in a pocket or bag or out of them). This includes the use of light or sound levels for identifing adequate positions and confining the sampling to the moments that the mobile is in those positions. For instance, the data collection technique in Miluzzo et al. [2008] takes photos when the user touches a key on the phone

[1]Arbiter pictures from the environment can be used for user localization [Ofstad et al. 2008]; information about the orientation of the phone can determine whether the colors belong to the ceiling or floor.

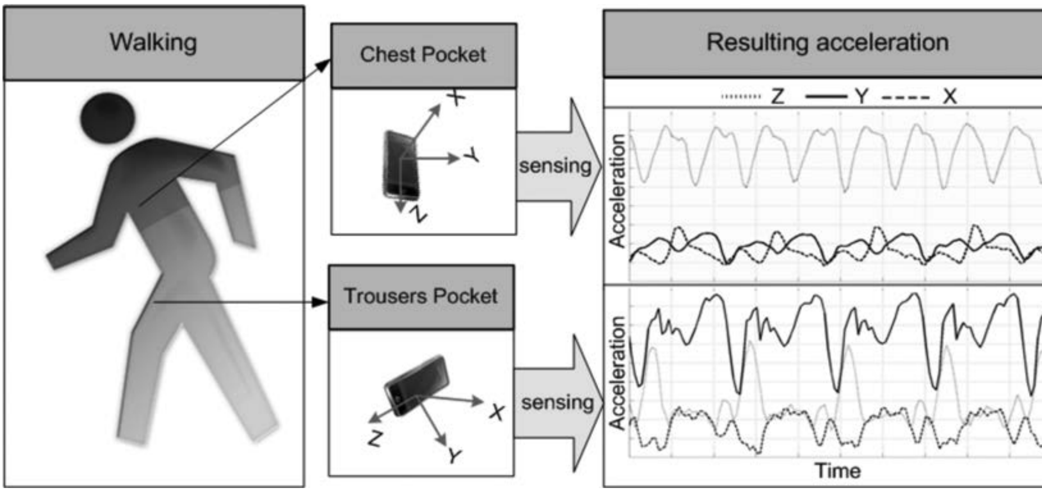

Fig. 3. Variation in accelerometer orientation and position affects the measured acceleration pattern. Magnitude, frequency of the components, and the axis of major components differ based upon the sensors' relative position and orientation relative to the user.

or, in Azizyan et al. 2009, a photo is taken when the user is answering a phone call. In the SoundSense project [Lu et al. 2009] an admission control stage is designed which discards the samples with unacceptable quality due to an inappropriate phone context. Many studies have investigated the adverse effects of misplacement and disorientation of the inertial sensors on the recognition and classification process (e.g., [Mathie et al. 2004; Gyorbiro et al. 2009; Olguin and Pentland 2006]). Figure 3 shows how the variation in position and orientation of a device affects the sensed acceleration data while walking. For inertial sensors, the problem of misplacement is usually solved by providing a position detection stage before preparing the data for feature extraction or classification or else by training the classification algorithms for all possible positions of the device. In some studies, the users are even asked to keep their mobile devices in a particular position. This makes the resolution of disorientation errors easier to accomplish and requires minimal involvement from the user. In case of device orientation, the samples are combined to an orientation-independent form, or data from the magnetometer and accelerometer sensors are processed to perceive the orientation of a device. An introduction to a variety of these techniques is presented in the following section.

3.1.1. The Effects of Device Position. The dependency of magnitude and the frequency of measured acceleration on the position of the accelerometer on the user's body has been highlighted in Section 2.1. Different studies have attempted to address the effects of inertial sensor positioning. Researchers in the area of BSNs, for example, have looked at the placement of sensors from the perspective of wearability and user convenience by letting the user decide the body position of the sensors [Kunze et al. 2005]. These related methods and algorithms can be classified into the following three cases.

One set of methods trains the classification algorithm on all possible positions. The algorithm subsequently tries to detect the context directly, regardless of a mobile's position. These methods usually require large databases and are less accurate when compared to other models. However, classification is achieved more quickly. Calibration is not required with this method because all possible positions are predefined. In other words, the observation is assumed to always be performed in one of the predefined references that the classifier is trained for. For example, in Lester et al. [2006], training the device with generalized data from different locations has shown that a reasonable accuracy can be achieved regardless of the phone's position. However, the accuracy of the model increases significantly with the number of individual training datasets. Lester et al. concluded that if the appropriate data from different individuals with different characteristics is available, the model can be used as a generalized model. Another example of such methods is presented in Brezmes et al. [2009]. Here, the classification method is trained based upon the user's preferred mobile position. The model can then distinguish between different user activities.

The second set of methods first infers the device position and then calibrates the data and features based upon the detected position for use in the classification algorithm. In contrast to other methods, the specific characteristics of the pattern of movement during certain activities are used for inferring the device's location. These methods rely upon extracting a number of features, which can be used to differentiate between different positions of the device during a certain activity. Although these methods are more efficient in memory consumption and give better accuracy during classification, they are usually more computationally expensive and require more time for recognition. The techniques are typically limited to a set of particular activities and corresponding positions of the mobile phone and are very susceptible to misdetection if they fail to determine a position correctly. For instance, if the positions of the mobile phone change during an activity, or a particular activity is not performed in a specific amount of time, the system is unable to calibrate itself. In Kunze et al. [2005] use the accelerometer signals during walking for recognizing the device position. Walking has been chosen as the example activity because it can be detected regardless of accelerometer position and orientation and is a very frequent activity in everyday life. Examining several positions on the body, such as wrist, head, trousers pocket, or chest pocket, this technique is reported to provide very high classification accuracy. Nevertheless, each segment takes more than three minutes to prepare for activity recognition. Kunze and Lukowicz [2007] have extended the previous work to sense the device position through a range of activities using accelerometer signal features (such as standard deviation, zero crossing, mean of the norm of the acceleration vector minus gravitational pull, and the absolute value of the number of peaks along the three axes). Kawahara et al. [2007] have exploited the unique behavior of accelerometer signals in multiple situations to infer the phone's position. Their threshold-based device position and activity recognition model is reported as giving a very high accuracy. In order to determine the position of the phone when worn in a chest pocket, accelerometer patterns caused by a person stooping forward in the chair were used. For the trouser pocket position, fluctuations of the tilt angle during walking were used. Furthermore, the variance of the signals provided sufficient indication to determine when the phone was not with the user. Finally, Miluzzo et al. [2008] have exploited vocal signals from a mobile phone's microphone to extract a set of required features to estimate the device position. Here, a sophisticated classification algorithm (a Gaussian mixture model with twenty components) is adopted to classify the position of the device as in or out of the pocket.

The third set of methods considers a fixed position for the sensing device in order to avoid an arduous calibration process. These methods give better computational efficiency and accuracy than previous methods, at the cost of losing generic applicability of the system. In order to find a proper position for such techniques, a number of positions have been proposed with different perspectives.

A review of the related literature in activity recognition with accelerometers suggests suitable positions near the center of gravity (COG) of the subject (e.g., [Mayagoitia et al. 2002; Sekine et al. 2002; Evans et al. 1991]). Murray [1967] shows that the applied force near the COG of the human body while walking is almost deterministic and undisturbed by individual characteristics.[2] The human center of gravity, also referred to as the body's center of mass, is located within the pelvic region when standing [Mathie et al. 2004]. COG is depicted in Figure 3.

Recent studies [Kawahara et al. 2007; Ichikawa et al. 2005] have identified the bag, chest, and trousers pockets as the most common locations where a user would typically carry a mobile phone during the daytime. Ichikawa et al. [2005] reported that women are more inclined to using bags, where men typically place their phones in their trouser pockets. However, the closeness of trouser pockets to a human's COG has made it a more attractive place for activity recognition tasks based on the inbuilt sensors (e.g., [Bieber et al. 2009; Kwapisz et al. 2010; Ofstad et al. 2008]). For example, Bao and Intille [2004] have investigated the effect of sensor position on mobile-centric activity recognition and suggest that positions near the hips are ideal. Inspired by Bao's findings, Miluzzo et al.'s study about different aspects of a mobile user's behavior Miluzzo et al. [2008] has encouraged participants to place their mobile phones in their front or back trouser pockets.

3.1.2. The Effects of Device Orientation. Not only the position, but also the orientation of the sensors, has an impact on measurements of the magnetometer and the inertial sensors along their sensitive axes. In other words, considering the same user context and position of device, the values that are sensed on a sensitive axis of a sensor would not be repeated unless the same orientation is used. Consequently, a major challenge of mobile phone-based sensing systems is the effect of frequent change in orientation of the mobile phone during everyday phone use and transport.

One common solution in overcoming the problems caused by disorientation is to transform the measured data into a scalar value and consider only the magnitude of the samples, that is, omit the directional data (e.g., [Gyorbiro et al. 2009; Yang 2009; Santos et al. 2010; Brezmes et al. 2009; Kwapisz et al. 2011; Fleury et al. 2010]). However, such techniques discard the valuable information that sensing in multiple dimensions could provide. Rather, some studies have developed some calibration techniques to have higherdimensional data, while trying to avoid the errors caused by disorientation. Particularly, for activity recognition, information should ideally be known in terms of a coordinate system oriented with respect to the user's body and aligned to his forward motion [Mizell 2003]. Figure 4 depicts the user body coordinate system. The user coordinate axes are denoted as V (for vertical vector), F (for the user forward directional vector), and S (for the user side direction vector) which is the cross product of F and V. One of the key parameters in detecting a mobile phone's orientation is gravitational acceleration, which is parallel to the V direction of the user's coordinate system (see Figure 4) with a constant magnitude. In particular, as indicated in Section 2.1.1, the main variability of acceleration measurements in daily activities is in a user's vertical direction. As a simple and computationally efficient model, averaging accelerometer samples in a window of a few seconds provides a proper estimation of the gravitational vector [Mizell 2003]. A better approximation for gravitational acceleration is obtained by averaging the accelerometer samples at the moments when their variation in the sample window is almost zero [Kunze et al. 2009]. Another approach determines the gravity acceleration by separating out the body movement acceleration [Allen et al. 2006] by means of a low-pass filter with a cut-off frequency of approximately 0.25 Hz from the overall measured acceleration signal. Luinge et al. [1999] and Kourogi and

[2]This fact has also been used for reducing computational cost for activity recognition since no learning algorithm for absorbing individual characteristics is required any more (e.g., [Kourogi and Kurata 2003]).

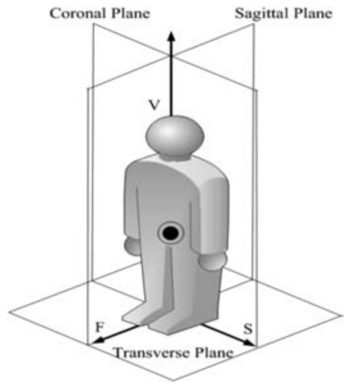

Fig. 4. Demonstration of a body's coordination system and rotation planes. The intersection of the planes shows the position of CoG.

Kuratta [2003] use gyroscope measurements for determining the device's orientation. The device's orientation is calculated by passing accelerometer and gyroscopes measurement values through a Kalman filter. According to Zhang et al. [2008], processing gyroscope signals typically requires a large number of sine/cosine and coordinate transform operations, which puts a heavy computational burden on the processor, making it less suitable for mobile computing environments. Consequently, Zhang et al. concluded that if a task could be identified only by accelerometers, the use of gyros should be avoided. Detecting the gravitational vector in turn gives an estimation of the vertical component of user motion (parallel to gravity) and the magnitude of the resultant of horizontal components. However, the direction of the horizontal components remains undefined. Taking into consideration only the magnitude of the horizontal and vertical components as a two-dimensional measurement has been shown to provide good accuracy for activity recognition on mobile phones [Yang 2009]. Despite the success with two-dimensional measurements, some studies have even developed techniques that provide the direction of the axes (F and S) in the horizontal plane. For example, the application of principal component analysis (PCA) to accelerometer signals has been proposed [Kunze et al. 2009] in order to determine the forward direction of users (i.e., F in Figure 4) in the horizontal plane. The resulting accuracy is reported to be comparable to those approaches using GPS. The PCA method, which uses only the identities of multiplication and addition, is considered a computationally efficient method and has been successfully implemented in this study in a mobile device (a Nokia 810). A novel semi-analytical approach has been recently presented in Hoseinitabatabaei et al. [2011], where a dynamic model of the movement of the body segments corresponding to the position of the device is used for recognizing the coordinates of the user. The model has been shown to outperform the conventional PCA- and GPS-based approaches. Combining these techniques with the vertical direction identification provides a calibration method for transferring the observations into the user body coordinate system.

To summarize, using the mobile phone as a sensing platform requires detection and compensation of disorientation and misplacement, especially when inertial sensors and magnetometers are involved. A variety of techniques were introduced in this section to tackle these problems. Table I summarizes the mentioned techniques alongside appropriate examples of available approaches. Having all the sensor data from the predefined references after the calibration process, the next step in preprocessing is to extract features from the calibrated data.

Table I. Calibration Process: Aspects and Available Techniques

Calibration					
Position			Orientation		
Training on All Possible Positions	Detecting Device Position	Using a Particular Position	Invariant (One Dimension)	Two Dimensions	Three Dimensions
[Lester et al. 2006] [Brezmes et al. 2009]	[Vahdatpour et al. 2011] [Kunze and Lukowicz 2007] [Kawahara et al. 2007]	[Miluzzo et al. 2008] [Mayagoitia et al. 2002] [Sekine et al. 2002] [Evans et al. 1991]	[Santos et al. 2010] [Brezmes et al. 2009] [Kwapisz et al. 2010]	[Yang 2009] [Lu et al. 2010]	[Kunze et al. 2009] [Hoseinitabatabaei et al. 2011]

3.2. Feature Extraction

Feature extraction is the process of distilling the raw sensor data and converting it into more computationally-efficient and lower-dimensional forms called *features*. Typically, the raw sensor data is first segmented into several windows, and features are extracted from a window of samples. It should be noted that the window size is an important parameter that affects both the computation and the power consumption of sensing algorithms [Himber et al. 2001] and is also required for minimizing jitter [Santos et al. 2010]. However, a detailed analysis of the effect of window sizes is beyond the scope of this article.

The generated features represent the gist of information from a window of raw samples. Features from sensor readings are often used as inputs into the classification algorithms (Section 4) for recognizing user context. In this section, a variety of feature-generation techniques are introduced within a number of different subcategories. First, heuristic features refer to features that are derived from a fundamental and often intuitive understanding of how a specific aspect of a user's context would be determined from a sensor's readings. We have described user context as a physical activity, environment, and/or social interaction. Other subcategories of features are time and frequency domain. In contrast to heuristic features, time and frequency domain features are simply used to characterize the information within the time-varying signal and are not typically related to specific aspects of context. Compared to the time domain, the frequency domain features require a further preprocessing stage of transferring sensed data from the time domain to the frequency domain. Due to this added process, generating the frequency domain features is regarded as more computationally demanding than time domain features [Miluzzo et al. 2008; Gyorbiro et al. 2009]. However, very fast and efficient domain conversions are now achievable with different computationally-efficient versions of fast Fourier transforms (FFT), such as the Fastest Fourier Transform in the West (FFTW) [Frigo 1999].

There are a large number of features that can be generated through different mathematical and statistical procedures. This is particularly true when offline processing is available in backend servers with no limitation in processing time, memory, and energy consumption. However, for processing data on mobile phones, these limitations must be carefully considered. Accordingly, we focus our discussion on features for user context recognition that have been successfully examined in miniaturized processors used in mobile phones or PDAs.

Selecting the most informative feature and sensors is critical to reducing power consumption, learning, and classification problems [Choudhury et al. 2008]. For that reason, a sensing system should ideally be able to dynamically select between different features and sensors in different situations. Meanwhile, the level of information that is conveyed by the generated features from a particular sensor is closely related to the desired context. For instance, while determining the standard deviation from a window of accelerometer samples can provide a substantial amount of information about a user's physical activity, it would be less useful for determining user social interactions. Therefore, we have further classified features based on their main context of application, namely user physical activity, social interactions, and environment.

3.2.1. Features Used in Physical Activity Detection. Methodologies from the realm of mobile-centric sensing have taken advantage of the ubiquitous presence of mobile devices in order to observe fragments of user physical activities in unfettered conditions. In the case of young adults and children, the main fragments can be categorized into a few groups. Based on the reported results of a comprehensive survey [Bieber et al. 2009], the most commonly performed activities during a day are lying down (ca. 9 hours), standing (ca. 5 hours), sitting (ca. 9 hours), and being active, for example, walking, running, etc. (ca. 1 hour). In an effort to observe at least a subset of these fragments, many studies have exploited mobile embedded sensors for activity recognition. The main contributing sensors for capturing these contexts are inertial and positioning sensors. While the inertial sensors can discriminate among a variety of daily physical activities, the position-based method can distinguish between different modes of movement. Accelerometers are considered to provide the most discriminative information for activity recognition [Choudhury et al. 2008; Lester et al. 2006]. Accelerometers have been extensively utilized for determining a variety of activities, such as walking, running, standing, or sitting (e.g., [Miluzzo et al. 2008; Yang 2009; Ravi et al. 2005; Azizyan et al. 2009], respectively) and sometimes also climbing (e.g., [Kwapisz et al. 2010]), cycling [Bieber et al. 2009] or driving [Ermes et al. 2008]. Diverse studies concerning the accelerometer features in different activity recognition systems demonstrate that simple time-domain-based features are usually adequate for detecting the majority of demanding activities (e.g., [Allen et al. 2006]). Despite the remarkable potential for detecting user rotational movements, magnetometer samples have been less frequently used for mobile-centric activity recognition to date (e.g., [Choudhury et al. 2008]). In this section, the main features generated from different mobile embedded sensors are presented.

Time-Domain Features. Mean and *standard deviation* are the most commonly used time-domain features for accelerometer signals [Miluzzo et al. 2008; Ermes et al. 2008; Santos et al. 2009; Kunze and Lukowicz 2007; Sashima et al. 2008]. The signal average is often taken so as to differentiate between different body postures of a person. In such cases, the deviation from the mean can distinguish standing from sitting [Yang 2009, Miluzzo et al. 2008]. The signal variance is also used as a natural choice for estimating the intensity of activity. For example, Ermes et al. [2008] calculated the variance of samples in order to distinguish running from walking and averaged the variance over all the axes of accelerometer data in order to identify the standing state [Ofstad et al. 2008]. Yang [2009] has also used the mean and variance of horizontal and vertical acceleration for activity recognition. Another common feature is the *number of peaks* per unit of time along the three axes of the accelerometer for distinguishing between walking from running [Miluzzo et al. 2008; Kunze and Lukowicz 2007].

In another approach, researchers have used the *intensity* of the signal as a feature, claiming that it is directly proportional to the acceleration [Gyorbiro et al. 2009]. The intensity is calculated as the sum of the numerical derivative of a window of samples, normalized to the length of the window. The derivative of the acceleration samples in calculating intensity reflects the volatility of the samples during the performed action.

Apart from these preceding accelerometer-based features, logging the pattern of user locations over time is often sufficient to detect the user's activity motion. As a result, all the sensing systems that are introduced for localization techniques, in principle, are able to provide such information about the user. However, the level of recognition varies from very abstract states such as moving or stationary mode to finer-grained levels, such as walking, driving, and running, based on the accuracy of the implemented technique. Some examples of such systems for mobile phone-centric sensing are now provided.

The GSM signals received on mobile phones have been a conventional source for inferring different states of user motion (e.g., [Sohn et al. 2006; Anderson and Muller 2006b]). By means of a different features, such as *signal strength* and *cell tower fluctuations*, user movement activity is estimated within a time window of few tens of seconds. In Anderson and Muller [2006b] and Anderson et al. [2007] use the change in the number of unique LACs, along with the fluctuation of signal strength and the rate of change of cells, for identifying different modes of movement. GPS is also widely used for detecting movement. Miluzzo et al. [2008] use the *GPS positioning information* over time for inferring users' mode of movement, such as being in a vehicle, running, or stationary, by estimating their speed. Since activity recognition with localization techniques requires a comparison of several subsequent locations of the user, these techniques typically require a greater amount of time to determine the state of the user than systems that take advantage of inertial sensors.

Frequency-Domain Features. Due to the computationally-efficient and sufficiently-informative features that can be generated in the time domain, converting sensor data into the frequency domain has been less popular in mobile phone-centric sensing. In Santos et al. [2010], FFT is performed on a window of accelerometer samples, and the amplitude and frequencies within the range from 0.5 Hz to 2 Hz are summed. The resulting feature (which corresponds to the energy of movement) is compared to a predefined threshold in order to distinguish fast movements from regular ones. In Ermes et al. [2008], the *peak frequency* of the power spectral density of the accelerometer signal served as a clue for detecting cyclic activities, such as cycling, walking, and running.

Heuristic Features. In the absence of motion, the accelerometer samples are equal to the cosine of the angle between the gravitational acceleration and the sensitive axis. Similarly, a magnetometer is able to detect the angles between geomagnetic fields and its sensitive axis. The fact that different activities change these angles in different ways has attracted the interest of researchers to use these features for activity recognition. Examples include the use of angles that are directly calculated from accelerometer measurements (e.g., [Kawahara et al. 2007]), magnetometer measurements (e.g., [Fleury et al. 2009]), or even the rate of change of gyroscope measurements (e.g., [Lee and Mase 2002]).

3.2.2. Features Used for Detecting Social Interactions. Perceiving social signals with mobile phones to attain insight into one's daily social interactions has gained the attention of a number of researchers. Social signals refer to the nonverbal behaviors that represent the expression of a person's attitude toward a social situation and interplay [Viniciarelli et al. 2009]. For an extensive overview on social signal processing, the reader is referred to Viniciarelli et al. [2009]. Amongst the different features that have been used for mobile-centric detection of social interactions, the detection of social proximity has been given most significance, as the presence of other people in the proximity of a user is considered a necessary prerequisite for having a social interaction.

Time-Domain Features. In order to determine the presence of a social interaction as the first step for understanding social interactions, a number of techniques have been proposed. The SoundSense project [Lu et al. 2009] used Zero Crossing Rate (ZCR) and low energy frame rate (defined as the number of frames with an RMS value less than 50% of the mean of an entire window) for distinguishing human voice (presence of conversation) from music and ambient noise on a mobile sensing platform. Here, ZCR or number of zero crossing within a time frame can determine the human voice from music and ambient noise [Saunders 1996]. Calculating the low energy frame rate is also relevant since human conversations have more moments of silence than music and ambient noise [Saunders 1996].

The physical and nonverbal behavior of individuals convey a significant amount of information about their behavior in social interactions. A study by Viniciarelli et al. [2009] has identified the most important features of vocal and nonverbal behavior as voice *quality*, *turn talking*, and *silence/pauses* during speaking. These features can be extracted with a simple microphone, without directly analyzing the user's speech. Such information is used in persuasive applications (e.g., a personal tutor) for detecting the user's role in different interactions and by providing proper feedback [Pentland 2009]. For instance, microphones are used in *sociometer* badges (e.g., [Olguin and Pentland 2008; Kim et al. 2008]) in order to detect social roles, the dominance in conversations, and the level of excitement and interest. Integration of these sociometer badges with mobile phones allows for direct feedback to the mobile phone user. The samples obtained from the accelerometers are also used to understand user social interactions. Kim et al. [2008] purpose *average of body movements* within a fixed unit of time during a conversation to help analyze of behavior (e.g., the level of involvement) during social interactions.

Frequency-Domain Features. Converting the microphone samples into the frequency domain for extracting features has been widely used for determining the presence of a social interaction. For instance, Miluzzo et al. [2008] have made use of the variance and the mean of a discrete Fourier transforms (DFT) of the recorded signal from a mobile phone microphone in order to distinguish conversation moments from ambient noise. Lu et al. [2009] have introduced and implemented a number of frequency-domain features for distinguishing the human voice from both music and the ambient noise on a mobile phone device. These features are described in the following.

—*Spectral Flux (SF)* is defined as a vector of 2-norm of frame-to-frame spectral amplitude difference [Scheirer and Slaney 1997]. SF has a different shape for typical music and voice signals, as music usually has less SF.
—*Spectral Roll-off Frequency (SRF)* is another feature, calculated as the 95th percentile of power distribution [Scheirer and Slaney 1997]. A larger number of highfrequency components in music compared to human voice leads to higher SRF.
—*Spectral Centroid (SC)* is defined as the balancing point of a spectral power distribution [Scheirer and Slaney 1997]. The use of SC relies on the difference of the spectral power distribution between the human voice and music.
—*Normalized Weighted Phase Deviation* [Dixon 2006] is determined by weighting the phase deviation of frequency bins in the spectrum by their magnitude. Ambient sound and music have less phase deviation than the human voice.
—*Relative Spectral Entropy (RSE)* is simply the KL (Kullback-Liebker) divergence between the current spectrum and the local mean spectrum [Basu 2003]. It is calculated from sound signals in order to distinguish human speech from other sounds.

Heuristic Features. Bluetooth scanning is the most popular technique for detecting social interactions. This makes use of periodic invocations of the Bluetooth device discovery function in order to determine the devices (and thus users of those devices) in proximity to the user. The presence of another user in proximity is considered as a potential social interaction. The technique exploits the uniqueness of the BTID MAC identifier, which is transmitted by mobile phones together with Bluetooth-personal area network capabilities when queried. Miluzzo et al. [2008] compared the logged BTID with a database of MAC addresses in order to infer whether a user is near his or her friends. The proximity information is further used for identifying persons in one's vicinity, in order to subsequently establish correlations between people with the same application and to calculate social status metrics, for example, popularity. A case study by Eagle and Pentland [2006] of the social interactions of students utilizing the logging of Bluetooth proximity has reported that there is a significant correlation between social interactions and the number of logged BTIDs when senior students were studied. However, for new incoming students, the correlation was not significant. An example of an application relying on such observations is the BlueAware platform [Eagle and Pentland 2005], in which the discovered BTIDs of neighbouring mobile devices are time-stamped and an reported to a backend server. The collected data is then analyzed to extract patterns of social relations, thereby suggesting the networks of social relations. Another example is the Jabberwockies system [Paulos and Goodman 2004], which uses Bluetooth scanning by mobile phones to demonstrate the relationships between commuters who do not know each other well but do see each other daily at public places, such as bus stops and railway stations.

3.2.3. Features from Environmental Sensing. The user environment has been observed from a diversity of perspectives. Conventional approaches in the mobile opportunistic-sensing realm are mainly identifying the user environment from a set of predefined classes of locations. Types of location classes range from absolute geographical locations to semantic and logical locations. The most common techniques take advantage of absolute positioning of users from GPS (e.g., [Cho et al. 2007; Gaonkar et al. 2008]) or GSM signals e.g., [Eagle and Pentland 2006; Laasonen et al. 2004; Bhattacharya and Das 1999; Bar-Noy and Kessler 1993]) to infer the user's location and overlay it onto a map using a geographic information system (GIS). Inertial sensors, such as accelerometers and gyroscopes, are also used to detect the user's movement pattern in a known topology, that is, dead reckoning (e.g., [Blanke and Schiele 2008][3], Lee and Mase [2001]. Information about user direction is typically obtained from magnetometers (e.g., [Lee and Mase 2002]. Compared to the first two categories of user-context sensing, determining qualities of a user's environment is typically carried out using heuristic features. The reduced use of time and frequency domain features for determining environmental context such as location can, in part, be explained by the reliance on absolute positioning systems, which usually do not require an analysis over time (or frequency).

Time Domain Features. Probability density functions (PDF) of the locations of cell towers that the phone is near over specific periods of time have been used for inferring the user location [Eagle and Pentland 2006]. In Santos et al. [2010], a window of samples from sensors, such as sound, light, temperature, and humidity, have been averaged and mapped into a specific category using different thresholds. Each category corresponds to a specific location (e.g., indoor or outdoor). Patterns of the acceleration samples generated in different locations are also used as fingerprints of the locations for logical localization. For example, Ofstad et al. [2008] use the percentage of time that a user is in a standing state for localization (e.g., being in a coffee shop or shopping center), where the standing state is determined from the accelerometer samples.

Frequency Domain Features. Only very few environment-sensing approaches have utilized frequency-based features on mobile devices. A recent approach in Lu et al. [2009] has exploited a number of frequency domain features from signals of a mobile microphone, in order to distinguish between ambient noise from music and the identification of distinct sound events. One such feature is *bandwidth*, which in spite of its

[3]Relative positioning is used in contrast to absolute positioning systems, such as GPS- and UWB-based approaches.

conventional definition can be regarded as a measure of the flatness of aFFT spectrum. While ambient noise has a limited spectrum, music is typically spread across a wider range of frequencies. Other features worth mentioning are *Melfrequency cepstral coefficients (MFCC)*. MFCCs are compact representations of a spectral envelope of audio signals and mimic the human perception of pitch in their calculations [Lerch 2009]. Although MFCC feature extraction is a computationally demanding process, MFCCs have been used effectively on mobile phones for recognizing distinct ambient sound events in the user's environment.

Heuristic Features. Heuristic features, which are usually assigned to the characteristics of different locations, are used to provide a logical localization. Some examples of the recent approaches are provided in the following discussion. Various approaches for recognizing a user's environment make use of features from camera pictures, such as illumination (e.g., [Azizyan et al. 2009]) or the colors (e.g., [Ofstad et al. 2008; Miluzzo et al. 2008], and even sometimes the contents extracted from the picture (e.g., [Kansal and Zhao 2007]).

Microphone samples for sensing the ambient noise level are used for logical localization (e.g., [Miluzzo et al. 2008; Ofstad et al. 2008; Santos et al. 2010; Azizyan et al. 2009]. For instance, Aziziyan et al. [2009] have used the noise level as a location fingerprint, while Santos et al. have used the noise level captured on a mobile phone's microphone as a clue for indoor or outdoor location [Santos et al. 2010]. In another study, the noise level indicates whether the user is attending a party [Miluzzo et al. 2008]. There, the noise level is then combined with other data obtained from accelerometers and Bluetooth to give a better indication of the social context of a user.

The absolute position of a user, determined by GPS or cell ID and corresponding cell tower signals, is mapped to the nearest pre-determined positions indicating user location with segment labels [Anderson and Muller 2006b; Laasonen et al. 2004; Arikawa et al. 2007]. This feature may then be used for detecting user landmarks [Cho et al. 2007]. A similar method in Miluzzo et al. [2008] has estimated user location based on manually labeled traces of GPS. Another approach uses a static Bluetooth beacon [Eagle et al. 2009] or WiFi transmitters [Miluzzo et al. 2008] to detect the presence of a user in a predetermined location. Here the reception of signals from several transmitters, each with a particular MAC address, indicates the location of a user.

The *received signal strength* from different radio systems has also been widely used for user localization recognition (e.g., [Meeuwissen et al. 2007; Laasonen et al. 2004]). For instance, in Eagle et al. [2009] logs the GSM signal strength on mobile devices in order to determine which cell towers are in the vicinity and, consequently, the location of the device.

3.2.4. Summary. In this section, we have introduced techniques that have been successfully implemented on mobile phones for converting raw sensor data into a variety of features useful for user context recognition. Classifying the features into three subcategories of time domain, frequency domain, and heuristic features, the most relevant features for different aspects of user context have been presented. Conceptually, our discussion could have also included time-frequency-based features such as wavelets. As discussed by Iso and Ymazaki [2006], frequency- and time-domain-based features from sensor data have poorer time-frequency resolution than wavelet transformations and consequently are not able to identify localized wave data present in sensor data streams. The research community has taken advantage of these features in a variety of contextrecognition applications. Examples include wearable device context recognition [Kunze et al. 2005], detection of transitions between physical activities [Fleury et al. 2009], and classification of walking on level surface to climbing stairs for healthcare approaches [Sekine et al. 2000]. However, in the case of mobile phones, where, in contrast to wearable sensors, the computational resources are used concurrently for a variety of different tasks, the realization of these computationally-demanding features has been limited primarily to offline modes in the past (e.g., [Iso and Ymazaki 2006]). Recently, emerging powerful microprocessors for mobile phones now make the use of time-frequencybased features on mobile phones feasible. Consequently, the research community has started to move toward using such powerful features for different aspects of user context recognition. As one of the few existing examples, Wittke et al. [2009] used Harr-like features, computed similarly to coefficients in Harr wavelet transforms, for detecting user activities and device movements. However, work in this area is still very limited and has therefore not been prominent in our previous discussion.

Differing from the discussed signal-oriented features, model-based features have recently started to attract researchers in the wearable research community, providing a more reliable alternative for physical activity recognition. Here, the features are generated according to the model of the body when performing a certain action. These features may include several subactions, body posture, or relative position of user and objects. Being driven from a body model, these features are more robust and less susceptible to variability of performing different activities [Zinnen et al. 2009]. For example, Zinnen et al. [2009] used various motion primitives, such as moving the hand up or down or turning the hands, along with some postural features, such as relative orientation of the hand to gravity or relative position of the hands to each other, and location information for activity recognition. As a result, the model-based features provide more reliable results than conventional signal-oriented features. However, this approach is limited to situations where sensors have to be attached to the body segments of a user. Also, their heavy computational burden still remains an issue for implementation on mobile devices. Table II in the Appendix gives the type of features that have been used in a variety of different mobile-centric context recognition systems.

Features generated from sensor data are used in classification algorithms to identify the user context. In the next section, a variety of context inference techniques that have been implemented on mobile phones are described.

4. CONTEXT INFERENCE

Once the features are derived from sensor data, they are inserted into a classification algorithm (see Figure 1). Initially, each classifier requires a learning phase, where it learns the requisite patterns within the input features with each dimension of the desired user context. Once the learning phase is completed, the classification algorithm is able to assign an unknown window of data to a particular user's context class. Different classification algorithms are characterized with different degrees of complexity, starting from simple threshold-based algorithms to advanced models, such as neural networks (NN) and hidden Markov models (HMM). However, the classification methods that are implemented on handheld devices must be adapted to the computational limitations of microprocessors and available memory and respective energy constraints of the battery-powered devices. Moreover, in many cases when real-time feedback is required, the delay in the context inference process is a further limitation.

4.1. Learning Techniques

Based upon learning characteristics, classification techniques can be divided into *supervised* and *unsupervised* learning models. *Supervised learning* refers to learning through example algorithms, where data and its corresponding classes were presented during the learning process. Alternatively, in *unsupervised learning*, true examples are not given as solutions [Pietquin 2004]. Selecting either type of learning model affects the design of the labeling process, which is explained in Section 4.3. Normally, the aim of a learning technique is to minimize the *generalization error*. The generalization error refers to the expected error of the real testing data and not necessarily the training data. One major problem which arises during training (or learning) classification models that causes significant generalization errors is the *bias-variance* trade-off. According to Friedman [1997], the mean square of classification error (MSE) can be decomposed into three terms.

$$MSE = Noise^2 + Bias(f(x))^2 + Var(f(x)), \tag{1}$$

where x is the input feature vector and $f(x)$ is the estimation of the classification model for the class of x (where a particular class is of user contexts). In Eq. (1), *Noise* represents the irreducible error due to noise in the system. *Bias* is the error related to the selected method of learning (linear, quadratic, etc.), and the variance (*Var*) is the error related to the sensitivity of the classification model to the training set. In order to reduce the generalization or MSE error, both the variance and bias errors must be minimized. This is unfortunately not possible due to the natural bias-variance tradeoff. For example, while a learning model may suffer from underfitting problems (high bias error) due to training on very large datasets, it is also susceptible to overfitting (variance error) on a small training set and hence may lose its generality. This explains why sometimes simpler classifiers outperform more complex ones. Stable classifiers normally have high bias and low variance, while unstable classifiers have the reverse [Lotte et al. 2007]. While often constrained to simple classifiers, such as implementations on mobile phone devices and access to a limited dataset for the training process, researchers have been faced with variance error and unstable classifier problems.

A key to this issue is to have a stable classifier that scales to a larger number of users so as to improve the generalization of the training dataset. Particularly, when user-dependent parameters are learned (e.g., thresholds), the number of participants has a significant effect on the training procedure of models for general usage. A straightforward solution to this problem is to increase the number of participants during the collection of training data [Lester et al. 2006]. For instance, in Kwapisz et al. [2010], the model is generated and tested on 29 people, which gives it greater reliability when compared to similar studies with a small set of users, such as that of Yang [2009]. One of the main challenges of these approaches is estimating the number of required participants to have adequate independent training datasets. Despite involving wider ranges of people, researchers have tried to develop different, less time-consuming, and more efficient approaches. One example is *active learning,* where the initial labels from training data are used as a soft guess. By asking the user to check and even correct the misclassified results, the classification parameters are adapted to user characteristics during online learning (e.g., [Könönena et al. 2010; Brezmes et al. 2009]). *Community-guided learning (GCL)* [Peebles et al. 2010] is another available approach for generalizing classification methods. This work demonstrates that the classification accuracy of the available techniques can be improved by using crowd-sourced labeled data for training, while the probable mislabeling errors, that is, human errors, are addressed by using the data similarity. A combination of community-guided learning and active learning is proposed by Berchtold et al. [2010], where a service-based recognition architecture is used for recognizing user physical activities. Here, a Global Trainer Service on a backend server trains different combinations of classifier module sets from user community data, which are then personalized by another service called Personalized Trainer Service using user-annotated data. The personalization information is then transmitted as a bit vector to the device for personalization of the available classifier modules on the device. A different approach is to use features that do not change significantly among different users during the learning process [Kawahara et al. 2007].

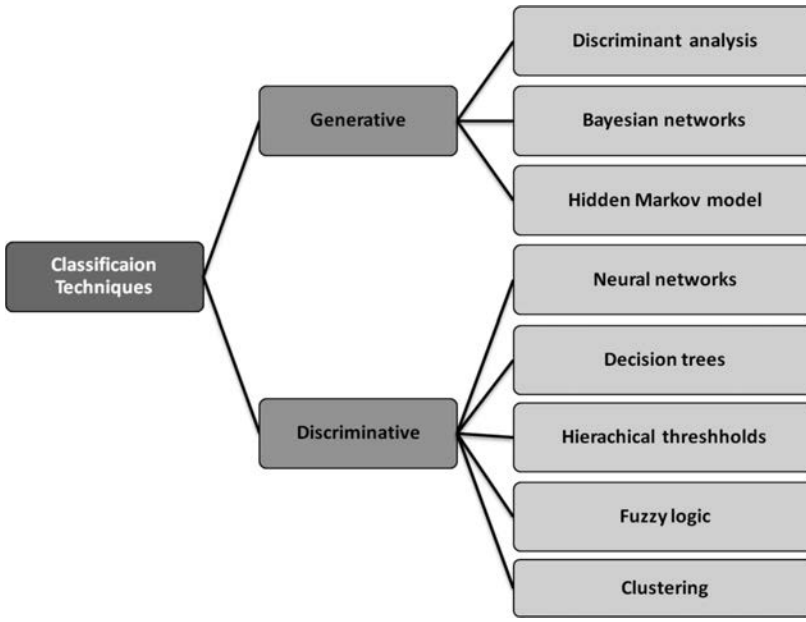

Fig. 5. A taxonomy of the classification techniques that have been successfully implemented for context recognition.

The learning techniques are determined according to the classification technique of choice. In the next section, different classification techniques used in mobile-centric applications are introduced.

4.2. Classification Techniques

As discussed by Ye [2004], almost all the classification algorithms are intended to solve an optimization problem. Based upon an optimization approach, they can be categorized as *discriminative* or *generative* algorithms.

The generative models assume a probabilistic pattern dependent on certain parameters, between data and classes, and specify a joint distribution over features and recognized classes. These can provide a direct model or a conditional distribution of data through *Bayes* rule. A generative classifier tries to estimate the underlying parameters and uses them to update the data classifications. Here *maximum likelihood* (ML), *mean posteriori*, or *maximum a posteriori* (MAP) techniques are used to perform parameter estimation. In the case of deterministic models, the only assumption made is that a well-defined distance and similarity measure exists between any pair of patterns. In other words, samples corresponding to one class may have a high similarity but are dissimilar to samples that belong to other classes, corresponding to a memory-based and nonparametric approach. Generative models have not been very popular due to their computational costs. In contrast, discriminative models have been widely implemented on mobile phone devices.

While many studies have used mobile phones only as a portable sensing system and then performed the data analysis and classification on backend servers, our emphasis for a mobile-centric sensing system is on classification techniques that have been implemented on mobile devices. Figure 5 shows a taxonomy of the algorithms that will now be presented, analyzing the recent approaches in developing classification algorithms on mobile phones.

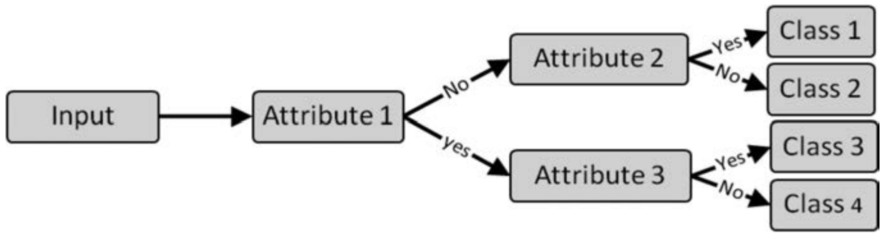

Fig. 6. Structure of a decision tree with three attributes which classify the input into four different classes.

4.2.1. Discriminative Models. A variety of discriminative models have been implemented on mobile devices. The most popular models include *decision trees*, *neural networks*, and *clustering techniques*. The major problem with many discriminative models is the susceptibility to overfitting (variance) [Deselaers et al. 2008] when creating rough boundaries between different classes of data during the training process. An introduction to the discriminative algorithms that have been successfully implemented on mobile devices is now presented. While discussing the different characteristics of classification algorithms, pertinent examples from mobile-centric sensing systems are provided.

4.2.1.1. Decision Tree. Typically, a decision tree consists of several nodes, branches, and leaves, where, during classification, each node examines an attribute. Each branch corresponds to an attribute value, and the leaves are classified context. Decision trees use rigorous algorithms that automate the process and create a compact set of rules [Webb 1999]. A sample for a decision tree which determines four classes is depicted in Figure 6.

Once the tree structure has been created, using a learning algorithm such as *ID3* (Iterative Dichotomiser 3), *C4.5*, or *J.48*, the process of classification with the decision tree is very fast. For example, the computation time required for a J.48 decision tree algorithm, used in user social-context recognition and in feature extraction (the mean standard deviation and a number of peaks in acceleration samples), together has been less than one second on a Nokia N95 [Miluzzo et al. 2008]. A comparison between ID3 and C4.5 [Santos et al. 2010] has shown that ID3 is superior to C4.5 on a Nokia N95 when classifying activities, such as walking, running, sitting, and standing, and logical location, such as inside or outside. Here again, relatively fast classification ($<.04$ s) and high accuracy have been achieved. Decision trees are one of the most popular methods, due to computational efficiency, especially when using trees of smaller scales. A comparison in Yang [2009] between different classifiers of a user's physical activity, using simplified features suitable for mobile applications, has shown that decision trees can obtain higher accuracy than naive Bayes or K-nearest neighbour approaches. Moreover, compared to threshold-based models, which are similar in concept, decision trees require less user intervention.

Implementing decision trees requires the consideration of several aspects: first, like many other algorithms, the learning process is time consuming. As a result, many studies perform offline training and only implement a final decision-tree classifier on mobile devices (e.g., [Kawahara et al. 2007; Santos et al. 2010; Miluzzo et al. 2008]). This limits the retraining process that may change the structure of the tree. Moreover, although decision trees with small sizes are computationally efficient and can be used effectively in real time [Maurer et al. 2006], increasing the tree's size can be computationally expensive since their evaluation is based on logical operations [Atallah et al. 2009]. Finally, decision trees are very prone to overfitting problems [Blum et al. 2006; Santos et al. 2009] and cannot be used for generic applications unless large datasets are available for training.

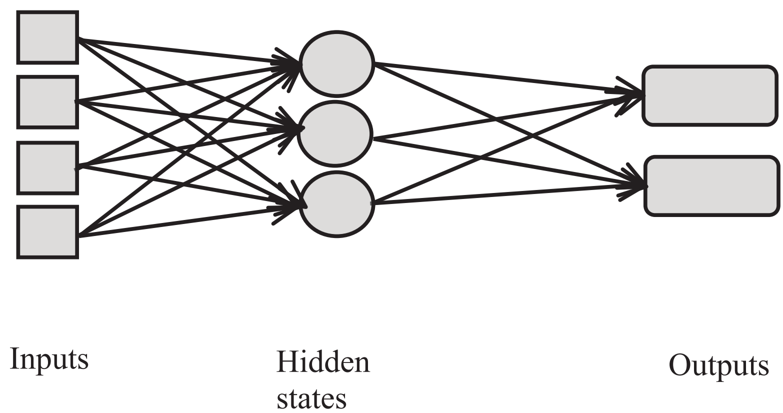

Fig. 7. Structure of feed-forward neural network with four inputs, three hidden states, and two classes of outputs.

4.2.1.2. Neural Networks. The work undertaken on artificial neural networks is motivated by complex, nonlinear, and parallel computation methodologies of the human brain. Neural networks use a connectionist approach to compute outputs through a network of inputs, hidden states, and possible outputs. Typically, neural networks can be divided into feed-forward networks, where signals can only move forward, and feedback networks, which also allow feedback loops in the network. The correct numbers of hidden neurons is found by training and evaluating the performance of classifiers with different numbers of hidden neurons. A feed-forward network with three hidden states is depicted in Figure 7.

Bruns et al. [2007] have successfully trained and implemented a two-layer neural network on a mobile device in order to recognize objects taken from a smartphone camera. Another example [Anderson et al. 2007] has implemented a neural network with eight hidden neurons (states) to map the pattern of signal strength fluctuations and changes in number of unique cell IDs to a user's state of activity.

In physical activity recognition, neural networks perform particularly well when only one activity needs to be detected [Gyorbiro et al. 2009]. Instead of using a large network for the recognition of various physical activities, Gyorbiro et al. have proposed a novel technique that allocates one neural network to each activity. Then, the network with the highest confidence determines the recognized activity. Similar to the decision trees, training neural networks is usually considered computationally expensive and consequently performed offline (e.g., [Gyorbiro et al. 2009; Anderson and Muller 2006b]). Therefore, neural networks are not suggested when the system is subject to frequent retraining.

4.2.1.3. Hierarchical Models. A hierarchy of thresholds has been used as a simple and computationally efficient model for mobile-centric applications (e.g., [Kawahara et al. 2007; Siewiorek et al. 2003]). Hierarchy models are very similar in principle to decision trees, with the exception that the training process is performed in supervised mode. For example, the E-coaching application [Kawahara et al. 2007] has implemented several thresholds based on the characteristics of different body movements in order to infer user activity and mobile device position. Although the thresholds are learned from empirical experiments, the variation of these thresholds between different subjects was found to be small enough so that they could be used in a generic solution. Similar to decision trees, the main weakness of this technique is a susceptibility to overfitting. Its

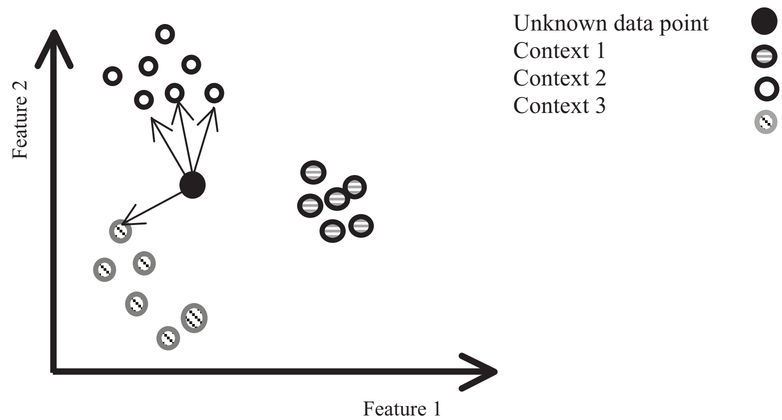

Fig. 8. KNN clustering: the input is the unknown data and its four nearest neighbours. The clustering is performed in a two-dimensional feature space.

dependency on user supervision during training (or retraining) is another constraint for the application of this method.

4.2.1.4. Fuzzy Logic. Similar to a human's understanding of a physical process, fuzzy logic is able to embed imprecise and approximate reasoning (instead of precise quantities that are used in computers) for solving complex problems [Ross 2004]. Fuzzy logic maps a set of inputs to one or more outputs with an assigned membership value or fuzzy truth via a set of if-then rules. Normally, the output with the maximum fuzzy truth is then taken as the result. Considering that the reasoning is based upon imprecise concepts, fuzzy logic seems more appropriate for real-world applications than conventional logical reasoning in hierarchical or decision trees [Preece et al. 2009]. In spite of this, only a limited number of studies have applied fuzzy logic in their classification problems. For example, Haykin [2009] used fuzzy logic for selecting the most probable state from outputs of a group of neural network classifiers for physical activity classification on a mobile device. A combination of decision trees and fuzzy logic has been used [Lee and Mase 2002] for indoor localization applications, where the fuzzy model is able to classify walking movements as slow, normal, or fast by defining several thresholds for acceleration and angular features.

4.2.1.5. Clustering. Despite the aforementioned issues with supervised learning algorithms[4] which require labeled data during training, some studies have used clustering as unsupervised learning algorithms for both classification (e.g., [Brezmes et al. 2009]) and calibration (e.g., [Anderson and Muller 2006b]). The clustering is described as an unsupervised classification of patterns (observation, data items, or feature vectors) into groups of clusters [Jain et al. 1999]. For an extensive discussion about different clustering techniques, refer to [Jain et al. 1999].

KNN Clustering. Naturally, our intuitive notion of a cluster is a group of entities in proximity to each other. In that sense, the nearest neighbour distance serves as a basis for clustering procedures for *Knearestneighbours* (KNN) algorithms. In KNN, unlabeled data is processed in multidimensional feature space containing all training data points corresponding to different contexts. The new data is labeled based upon its distance from a particular labeled data. Figure 8 represents a schematic of the KNN classification process.

[4]The learning process for artificial neural networks can be either supervised or unsupervised.

The activity recognition technique [Brezmes et al. 2009] used the *K-nearest approach*, which is trained based upon userpreferred mobile position and a specific set of activities. The data is classified based upon the Euclidian distance of present record towards predetermined data. The reported accuracy after full training was more than 70% for all activities. In another approach, the KNN classifier is used to classify the users' locations [Ofstad et al. 2008].

K-Means Clustering. Mirkin [2005] described the *K-means* algorithm as a major clustering technique which is fast and straightforward. In this technique, a multi-dimensional space of features is divided into *K* clusters through a recursive algorithm that finds the optimum position of cluster centroids. Although the K-means algorithm is fast and computationally efficient, it relies on saved data, and its implementation on mobile phones is limited by memory constraints. The K-means algorithm is also susceptible to local minimas, and attaining a global result may require several iterations of the algorithm. Due to these shortcomings, some studies (e.g., [Blum et al. 2006]) have deemed that the K-means algorithm is not a proper choice for classification on mobile phones. In the Shakara project [Anderson et al. 2007], the K-means algorithm is used as an unsupervised calibrating approach to learn the distribution pattern of the data which is used for quantizing the inputs of another classifier (HMM).

Another work reported in Yang [2009] uses mobile phone sensing for generating a user's physical activity diary. In this study, K-means clustering is used for smoothing out the classification results of a decision tree. Using K-means clustering, the magnitudes of the mean and standard deviations of accelerometer signals are divided into six clusters. The clustered data are then labeled for different classes of decision trees based upon the distances from their corresponding centroids. It is, however, unclear as to whether the algorithm has been actually implemented on a mobile phone.

4.2.2. Generative Models. Generative analysis, such as a hidden Markov model (HMM) or its hierarchical extensions, demonstrates significant potential for the classification of everyday activities. However, there is a significant challenge for porting resource-intensive HMMs into a mobile device. As a generative model which does not involve many mathematical calculations, discrete HMMs have widely been used for smoothing the classification results by finding the most probable output, considering one or a number of previous states [Wu et al. 2007; He et al. 2007]. For a detailed discussion of related issues, the reader is referred to [Atallah and Yang 2009]. The same resource requirement problem exists when conditional random fields (CRF) and dynamic Bayesian (DB) networks are used. Despite this issue, excellent classification results for offline implementation of CRF and DB have been reported (e.g., [Gyorbiro et al. 2009]). When computational resources are limited, the use of *Bayesian classifiers* (BN) represents a valid option for classification [Atallah et al. 2009].

Providing a probabilistic classification, generative techniques are more resilient to data variations compared to models with logical if-then rules, such as decision trees and hierarchical models. Some examples of the generative techniques that have been successfully implemented on mobile phones are presented in the following section.

4.2.2.1. Hidden Markov Model. Cappé et al. [2005] informally introduce hidden Markov models (HMM) as a Markov chain that is observed in noise. This Markov chain is often assumed to take a finite set of states which are not observable (hidden states). Each state is associated with a probability distribution, and state transitions are governed by a set of probabilities. Observations as another stochastic process are linked to the Markov chains, and an observation can be generated for each state. Most of the HMMs can be divided into two principally different classes of models: *left-to-right* and *ergodic* models [Cappé et al. 2005]. Figure 9(a) shows a left-to-right HMM, where the Markov chain starts in a particular state and, after a number of transitions,

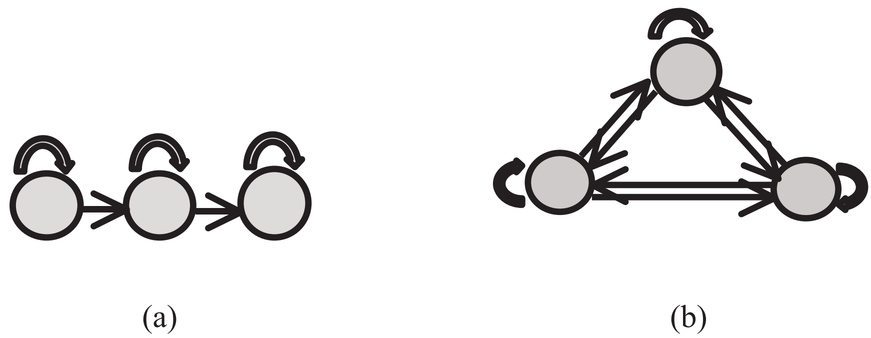

Fig. 9. Markov chain: (a) Structure of left-to-right HMM where transitions happens in the forward direction; (b) structure of the ergodic HMM where all possible transitions between states are allowed.

terminates in a final state. The transitions are limited to the forward direction (towards an end state). An ergodic HMM, in contrast, as shown in Figure 9(b), allows all possible transitions between states, and consequently it can produce an infinitely long sequence of outputs. When the distribution of observations is defined on finite spaces, the model is called a discrete HMM.

Anderson et al. [2007; Anderson and Muller 2006a], have implemented a discrete HMM model on mobile phones for recognizing user activity from GSM signals. Here, the observation data are based on signal strength fluctuation and cell fluctuations, which are mapped onto a set of 15 discrete observations. The hidden states describe the user's status (e.g., walking, driving, remaining stationary, etc.). The prediction is made based upon the sequence of five previous states. Markov models also have been used for smoothing out the classification results of other techniques, where the conditional dependency of the outputs is taken into account by training a Markov chain. For example, researchers in the SoundSense project have used a first-order Markov model to smooth the discrete classification results of a decision tree [Lu et al. 2009].

4.2.2.2. Bayesian Classifiers. As was mentioned earlier, generative models can produce conditional distributions of data through Bayes rule. Cakmaci and Coutaz [2002] have represented the Bayes rule formula for the context recognition as follows:

$$\mathrm{p(context|sensordata)} = \frac{\mathrm{p(sensordata|context)} * \mathrm{p(context)}}{\mathrm{p(sensordata)}}. \tag{2}$$

At this stage, different approaches have assumed different distributions for sensor data in each class. For example, *Naive Bayes* considers data points to be locally independent, while Gaussian discriminant analysis considers a Gaussian distribution in each class.

It should be noted that Bayesian classifiers are considered computationally efficient (containing only multiplication and additions) and can also be retrained by changing a few parameters instead of reprogramming the whole classifier (as it is the case for decision trees).

Discriminant Analysis. Gaussian discriminative analysis considers a multivariate distribution in n dimensions as follows.

$$P_k\left(x, \mu_k, \sum_k\right) = \left(\frac{1}{(2\pi)^{\frac{n}{2}} \left|\sum_k\right|^{\frac{1}{2}}} * \exp\left(\left(-\frac{1}{2}\right)(x-\mu_k)^T \sum^{-1}(x-\mu_k)\right)\right) \tag{3}$$

where, the subscript k indicates the class and μ is the mean vector $\mu \in R^n$ and $\sum$ is covariance matrix $\sum \in R^n * R^n$.

Blume et al. [2006], have implemented Gaussian discriminant analysis (also regarded as naïve Bayes with Gaussian distribution) to determine a user's speech, posture, and activity recognition.They claim the model is faster than HMM while providing comparable results, and is also immune to overfitting problems in contrast to the decision tree (C4.5) approach. The reported results show that the model has been able to distinguish between a majority of activities with high accuracy.

Note that since not all of the datasets can be approximated with a Gaussian distribution, it is sometimes required to extrapolate data with a statistical function, such as the kernel density estimation (KDE). For example, Ofstad et al. [2008] used KDE to implement a Bayesian classifier on a mobile device in order to infer the users' sitting and standing modes from the mobile's accelerometer data. As a result, very high classification accuracy has been achieved.

In contrast to Gaussian discriminant analysis, linear discriminant analysis considers the same covariance matrix (Σ) for all classes. As an example, discriminant analysis of audio samples for distinguishing human voice from ambient noise has been implemented in Miluzzo et al. [2008]. In this work, the targeted classes are learned over different samples of human voices (most of the energy between 0–4 kHz) with the mean and standard deviation as input features.

Bayesian Networks. Cho et al. [2007] have exploited modular Bayesian networks to recognize relevant or novel landmarks during movement in daily life and to visualize them as cartoon images. In order to implement a Bayesian network on a mobile device, a Bayesian network library for mobile devices called SMILE (Structural, Modeling Inference, and Learning Engine) is introduced. However, since monolithic models are susceptible to interference coming from large networks, an ensemble of multiple Bayesian networks specialized for each activity is proposed as modular Bayesian networks.

4.2.3. Classifiers Performance. When a classification algorithm is developed, it can be used for detecting a variety of aspects of user contexts. Here, the performance of a classifier in recognizing the demanded context determines the classifier of choice across other classifiers. In mobile sensing, typically the performance of the classifiers are studied in terms of their accuracy and computational complexity.

A comprehensive study of the performance of different classifiers of physical activities is presented in Preece et al., [2009]. According to Preece et al., an initial inspection of a variety of recent studies suggests that decision trees and neural networks are providing the highest level of classification accuracy. However, in some studies, the difference between classifier performances was not statistically significant and some classifiers, such as Bayesian networks, that were found to provide an acceptable performance for particular activities in one study have been reported as poor classifiers in another for the same activities. The same problem has been reported in a study about empirical evaluation of supervised learning algorithms by Caruana and Niculescu-Mizil [2006]. In this study, neural networks generally perform better in comparison with decision trees and naive Bayes. Here again, the results have shown significant variability between the performance of classification algorithms across problems and matrices.

This problem can be extended to other user contexts, including the detection of environment and user social interaction. Actually, the performance of classifiers is, to a great extent, affected by the context and the discriminative information in features that are used. Consequently, there is no classifier that performs optimally for all user context classification problems. Instead, one can always select a proper algorithm that provides the best classification accuracy with extracted features amongst all available classifications by evaluating techniques such as cross validation [Duda et al. 2000]. Although the required performance varies with different applications, there are some criteria that must be considered in assessing the classifiers.

The nature of the application is also a very important parameter in determining required accuracy and evaluating classifiers. In some applications, such as physical activity recognition, evaluating the overall prediction accuracy of the classifier is enough for selecting a proper classifier. Here, the distribution of the classes in the evaluating dataset is compared with outcomes of classifiers as a result of quadratic loss functions or cross entropy measurements or as a confusion matrix. An optimal choice is the one which minimizes the loss or off-diagonal figures in the confusion matrix. However, in another type of application, such as fall detection for elderly people, the cost of error in detecting the fall (TRUE Negative (TN)) is by far greater than the cost of error in detecting the normal movements (False Negative). This means that the priority is with the classifications that minimize the TN. Cost-sensitive classification or cost-sensitive learning techniques are different techniques for incorporating the misclassification cost into classifiers. Through the former technique. the costs are ignored during learning and are then applied to predictions as a cost matrix, while the latter considers the costs during the learning procedure. The required cost matrix in the first technique and the weights (costs) in the second are defined according to the application. Other widely-used techniques for accessing the performance of classifiers when costs are taken into account are to visualize them with receiver operating characteristic (ROC), cost curves, or calculating the F-measure score.

When the model is aimed to be developed on mobile devices with limited computational resources, it is very important to minimize the computational complexity of the algorithm. One common method for choosing between different models is to penalize the model complexity and minimize the following expression.

$$-2\mathrm{logL} + \mathrm{P}, \tag{4}$$

where L is the maximum likelihood and P is the penalty for complexity. An example of proposed form for P is Bayesian information criterion (BIC), where P is equal to$m \log n$, where m is the number of estimated parameters, and n is the sample size.

It is worth mentioning that the stated classification techniques are usually referred to as base-level classifiers. In addition to the base-level classifiers, meta and hybrid classifiers are also widely used. According to Ravi et al. [2005], meta classifiers can be classified into voting, stacking, and cascading types. Ravi et al. claim that the performance of base-level classifiers for activity recognition can be improved by using plurality voting techniques. However, the real-time implementation of these techniques remains an open research issue.

4.3. Labeling

Until recently, supervised learning techniques have typically been the algorithms of choice in building mobile inference systems [Lane et al. 2010]. Supervised learning requires all the possible classes of input data to be labeled before training. There are various ways to perform labeling on training data. Approaches have been developed by collecting user diaries or by making video tapes of them (e.g., [Fleury et al. 2010]), personal online labeling while data are gathered for learning (e.g., [Kwapisz et al. 2010]), and performing routine activities at particular times (e.g., [Mathie et al. 2004]). In other cases, participants in experiments have been asked to repeat the particular activities in the lab (e.g., [Kawahara et al. 2007]).

Performing a routine set of activities is susceptible to inserting bias in the data, which may result in producing optimistic data and so degrade the classification technique in reality [Azizyan et al. 2009]. Due to the dependency on hand-labeled data for training classifiers, applications that exploit these techniques are typically constrained to a small set of predefined aspects of user context (classes). Accordingly, a more challenging approach is to perform online learning and labeling in order to scale the available classes to a larger number of distinctive classes as required. Current efforts for labeling the new events that have not been covered during initial training utilize the user's intelligence when an unknown context is recognized. For example, Lu et al. [2009] purpose a hybrid approach for supervised and unsupervised learning where, after failing to recognizing the data by the supervised model, the unsupervised technique is used to learn a set of unlabeled classes that are frequently occurring. The user is then brought into the loop to provide a textual description (label) of the newly learned classes. A further example is Santos et al. [2010], in which the users are authorized to add (i.e., to label) their current contexts as a new context. Here, after user authentication, the device automatically learns the characteristics of the new context and retrains its classification algorithm.

5. COMPARISON OF DIFFERENT APPROACHES

Numerous options are available for selecting and integrating the aforementioned calibration, feature extraction, and classification techniques together to create an opportunistic sensing and context recognition system on mobile phones. However, careful selection of the functionalities and algorithms can both fulfil the application requirements, while minimizing the adverse effects on the user's phone experience. In this regard, a comparison of the suggested options and combinations of the available techniques is provided, aiming to provide further insights for researchers in this area.

Effective user context recognition on the mobile phone requires proper sensor and sampling frequency selection and sensor position and orientation calibration. It also requires noise reduction along with extracting informative features and selecting proper classification methods. Calibration can be done easily and with low amounts of computational cost. The calibration process is required to handle the daily life usage of mobile phones and can be divided into orientation and position calibration. The orientation calibration should ideally transfer sensor readings into a user's coordinate system. The necessary information can typically be derived from sensing the gravitational acceleration with an accelerometer, and processing those acceleration samples in a plane perpendicular to the direction of gravity. In order to determine the position of a mobile phone on the user's body, a variety of solutions have been proposed. Examples of these solutions are the collection of training data from all possible locations or even restricting applications to the most probable places where the device may be located. A popular example of the latter case is in a trouser pocket, as it presents a preferred location amongst men and is also in proximity to the humans' CoG. Moreover, when calibration is performed, the settings can be kept for a period of time and hence frequent updates are not required (e.g., until the user changes the position or orientation of the phone).

Simple time-domain statistical features, such as variance, meaning intensity, and number of peaks in a window of samples, seem to be essential inputs for inferring user physical activity. The most distinctive and informative features available for determining user social interactions are the user's proximity and vocal behavior. Finally, user environment can be characterized by combining absolute positioning data with heuristic features such as color, or typical user behavior, such as a location fingerprint.

Selecting a proper context recognition technique is one of the challenges that still requires further addressing. Before selecting a classification technique, an appropriate strategy for training and labeling is needed. Training the classifier may be performed either online or offline. Online training can provide a personalized training dataset and consequently higher classification results, while also imposing heavier computational burdens on the system. Alternatively, offline training is more computationally efficient but requires careful consideration about the generality of the training dataset in order to avoid overfitting problems. A hybrid combination may be achieved by providing a soft guess of the classes in offline training mode and then refining the misidentified classes with online training. This approach can be further enriched by community-guided data that is gathered and prepared in a backend server.

Once an online training mode is enabled, the system can be configured to learn the new classes of user context. However, still labeling the new context requires user intervention, which must be minimized in an opportunistic sensing system. Implementing unsupervised learning techniques to distinguish the most important unknown contexts, before involving the user, is proposed to mitigate this problem.

In the case of the classification techniques, an initial review of the introduced classification methods demonstrates that the decision trees and the neural networks provide satisfactory results for most of the applications. In small network (or tree) sizes, these can be easily trained and implemented on mobile devices. However, they are prone to overfitting problems. Developing hierarchical thresholds for hierarchical approaches is timeconsuming. However, similar to the decision trees, these can be executed with minimum power and computational cost and therefore are suitable for real-time applications. Neural networks also work well for complex pattern recognition, although usually the training stage is too burdensome to be performed on the mobile device. The Bayesian classifiers are simple to develop and can be executed rapidly and are also less susceptible to overfitting problems. However, they are based on weak assumptions about data distribution, and predictions are consequently not very accurate. Finally, HMM is a good choice for smoothing the prediction of other classifiers, including the effect of interdependency between different aspects (or classes) of a user's context. It should be noted that although many studies have compared different classification techniques for different purposes, there is no classifier that can optimally detect all aspects of a user's context.

Generally speaking, a two-level classification model, consisting of both a mobile device and a backend server, can fulfill the requirement of most applications. Inferring the context on the phone has been emphasized to provide a number of advantages [Miluzzo et al. 2008]. It presents resilience to cellular or WiFi dropouts and minimizes the data transmitted to the backend server, which in turn improves the system communication load efficiency. In addition, performing the context recognition process on the phone reduces the energy consumption of the phone and monetary cost by merging consecutive phone uploads. It also protects user privacy and the data integrity by keeping the raw data on the phone. Finally, it provides an opportunity for creating user-labeled contexts.

When a two-stage model is used, the inferred context or the learned parameters from user behavior can be provided to the backend servers for further processing. Especially in the case of real-time sensing applications, uploading the data to a backend server may help reduce the frequency of read and write events to the device. Note that writing to and reading from a data store can sometimes be the most time-consuming process of a mobile context recognition system [Santos et al. 2010]. The backend server can also provide the required connection (as a network) between other devices along with computational and storage support. Many studies have already exploited the more powerful computational capability of the backend server for further analysis of the data (e.g., [Miluzzo et al. 2008; Azizyan et al. 2009; Kanjo et al. 2009; Gaonkar et al. 2008]).

Finally, in order to control and minimize the power consumption of the sensing applications, a judicious selection of the different power-saving functions based upon application requirements, residual battery power, and a phone's current energy consumption profile is required. For example, when the locality of a user is required, one can take advantage of the energy-accuracy trade-off between different techniques, where energy consumption increases from GSM to WiFi-based localization and GPS schemes, while the accuracy decreases from GPS to WiFi and GSM methodologies [Gaonkar et al. 2008]. As another example, updating data on a backend server can ease the execution of burdensome tasks when an appropriate strategy controls the impact of the communication load and handset energy consumption (e.g., [Herrera et al. 2010]). A number of communication options are available for transferring the results to the backend server of a typical mobile phone device (e.g., Bluetooth, HTTP+3G, HTTP+ WiFi. The battery level of the device, the energy cost of the connections, and the available data rate and connection coverage are the parameters needed to determine the connection of choice. Some other suggestions are methods such as letting the user switch off the screen [Kanjo et al. 2009], selecting a proper sensor based upon the power demands and the required accuracy (e.g., [Gaonkar et al. 2008]), changing the sampling rate [Miluzzo et al. 2008], adapting the communication type (e.g., Bluetooth) to the user's activity [Crk et al. 2009], and offloading part of the data processing from the phone onto a backend server [Kanjo et al. 2009] to help reduce the power consumption. The proper application of such methods leads to developing a power-aware duty cycle for both sensing and uploading, while application responsiveness is not affected. Table II in the Appendix provides an overview of all the aforementioned aspects, from sensing to context recognition, for various applications.

6. CHALLENGES AND FUTURE OPPORTUNITIES

Technological advances in sensing, computation, and communications have turned mobile phones into pervasive observers. However, realizing the capabilities of such observers in real-life situations creates several challenges in terms of data acquisition and processing which need to be addressed. As mobile phones were not originally designed for sensing purposes, the main challenge is how to embed the required intelligence for pervasive observation without jeopardizing the phone experience. The following are some of the most significant challenges identified, and some recommendations are given.

6.1. Sensing

Despite the improvements in processing and storage capabilities, continuous sensing and context recognition can have an adverse effect on the responsiveness of other applications. Optimization of the sensing process to adaptively select sensor and sensing frequency on the phone would allow for a more efficient platform for pervasive observation. The other important challenge represents the limited control of sensors that is provided by device vendors in their SDKs (software development kit) and APIs (application programming interface). For example, it is currently difficult to establish a consistent sensing frequency that does not alter with CPU load. Effective programming for managing the sensing process can, to some extent, mitigate the problem. In the case of sensing frequency problems, for instance, some people have tried to interpolate the missing data caused by variations in sensing frequency [Bieber et al. 2009]. Another option is to rely on external hardware support that can be connected to mobile phones to overcome the design limitation of the manufactures (e.g., [Xsense 2011]

Finally, inspired by the fast growth of mobile-centric sensing applications, some researchers have determined that the sensing capabilities of neighbour devices can be utilized to improve the quality of the data [Dartmouth College 2010]. Such methods would help to gain access to sensing data from other devices when sensors are not available, or when the phone status is not appropriate for using them (e.g., not calibrated). However, it requires the devices to be able to establish a secure connection to other devices which may be using different APIs, thereby creating an open software issue [Lane et al. 2010].

6.2. Feature Selection

Feature selection is a decision-making process that connects raw sensor data to available featuregeneration techniques. Serving as a corridor between sensing and processing stages of a system architecture, an appropriate scheme of feature selection can substantially improve the energy and computational efficiency of the system. Performing a decent feature selection demands a precise consideration of a number of parameters.

Typically, it is preferable to use as few features as possible in mobile phone applications. This is for two reasons: first, the computational burden of feature extractions as the number of features increases, and second, the risk of obtaining suboptimal results due to classifier confusion when too many features have been used [Könönena et al. 2010]. While appropriate sensors are selected in the sensing stage, feature selection can confine the features to the most informative ones for a given sensor and the available classification technique.

In addition, the performance of different classifiers in terms of accuracy and their overall associated computational cost varies for a particular set of features. For instance, Könönen et al. [2010] have found that a relatively small difference between the accuracy of complex classification methods and a simple method can be achieved when features are properly selected. Moreover, there is a trade-off between the computational (and memory space) burden of the classification algorithms and the feature extraction procedure. The overall processing cost of implementing a complex algorithm can be comparable to a simple one, when simpler features are being used.

Finally, the extraction procedures of the features may overlap or depend upon each other. By avoiding the repetition in common processes, the overall computational and storage costs for feature generation could be reduced. For example, once FFT of the window of samples is calculated for deriving the spectral variance, many other features, such as energy and bandwidth, can be simply computed. The feature selection system must therefore be able to accurately consider the interdependency and overlap in the various combinations of features.

The current feature-selection approaches proposed for mobile-centric sensing (e.g., sequential forward/backward selection (SF/BS), sequential floating forward selection (SFFS) [Könönena et al. 2010], or the boosting-based technique in Choudhury et al. 2008]), although effective, are mainly from the realm of data mining to improve the classification results, and ignore a number of the aforementioned relations. Developing a technique targeting an optimal set of features within a mobile phone's computational constraints has remained a major challenge.

6.3. Labeling

Distinguishing and labeling different contexts forms another major challenge. In the real world, drawing boundaries between different aspects of user behavior is difficult. It is likely that people at home sometimes exhibit the same behavior as they do in their office or even perform different activities at the same time. The complex social behavior that people may exhibit in different conditions should be added to these facts. In this regard, providing a hierarchical context inference system that performs several levels of recognition with different time granularities and aspects of behavior appears to be essential for such systems in order to be useful in real-world situations. Another important shortcoming in current labeling techniques is their dependency on user intelligence when a new context is to be learned. Although, when managed properly, these techniques are considerably less intrusive, they still add a user bias into the data. Novel techniques built upon logical labeling from available clues in user context, such as *common sense* reasoning [Havasi et al. 2009], seem to improve the functionality of current systems to a large extent. One intuitive example is the unsupervised recognition of activities from *motifs,* as frequent repeated patterns, in time series of activities and labeling targeted activities (e.g., walking) according to their expected pattern (e.g., relative frequency of presence in a period of daily life time [Vahdatpour et al. 2009]).

6.4. Privacy

Another remaining challenge is determining how best to sense and exploit the data from the everyday lives of users, both locally on the device and globally on backend servers while maintaining user privacy.

Kapadia et al. [2009] envisioned some of the related security challenges in opportunistic sensing. They argue that the new characteristics of sensing architectures, including high mobility, opportunistic networking, strong but discontinuous connectivity, and relatively plentiful power in literally one hand, while dealing with very personal information on the other hand, has posed new challenges for information security. These challenges cannot be addressed with previous security solutions, such as cryptography and privacy-preserving anonymous aggregated data mining. The act of being sensed with other people in proximity, which is known as the *second hand smoke* problem [Lane et al. 2010], is also a challenge in mobile phone-based sensing. In addition, mobile phone devices are perceived as very personal items [Häkkilä and Chatfield 2005], and publication of the context information requires strict privacy and security considerations. Researchers have envisaged that privacy will remain a significant problem in mobile phone-based sensing for the time being [Lane et al. 2010] and that solving the privacy issue would be a significant step toward harnessing the potential of mobile-centric opportunistic sensing for real-world applications.

6.5. Identifying Potential Applications

The applications that could benefit from mobile phone-centric observations present exciting opportunities for further research. In the case of personalized applications, pervasive sensing technology can help the user to make more sophisticated decisions across a range of potential activities in order to select services and products considering the profile of the user and/or her goals.

More and more personalized applications based on opportunistic sensing are being introduced into mobile phones. A key question in this respect is what are the most likely upcoming applications in the next few years? The significant achievements of the wearable computing community in recent years provide some clues. We believe that with the continuous incorporation of new sensors into the mobile phones and advances in their computational resources, a wide range of these approaches will be available on mobile phones in the near future. The diversity of these applications will be very large, ranging from techniques for accurate relative indoor positioning (e.g., [Lee and Mase 2002]) to different approaches for enhancing the interaction of the user with smart environments, such as detecting and exchanging the user's physical states and orientation (e.g., [Ghiani and Paternò 2010]). The detection of the social and physiological context of a user, such as e-motions (refer to [Picard and Healey 1997]), social functioning and interactions, or even monitoring different health parameters (e.g., blood pressure, body temperature). Furthermore, the longitudinal data from user context that is gathered through these applications can be used to identify patterns or to profile the users for targeted advertizements or optimization of different service delivery to the user handsets. Recognizing the patterns and profiling different aspects of user behavior is a rapidly-growing research area and includes different perspectives of user behavior. One interesting existing example is the study of routines of physical activities of elderly people [Huynh et al. 2008], where the *topic modeling* technique is used to detect the patterns from longitudinal data that are generated via an activity recognition application. Another recent study profiles users according to the dynamics of their social networks [Candia et al. 2008] It is likely that in the near future, similar researches will extend the current boundaries to other aspects of user context and even to the study of the interconnections between different contexts.

In large-scale applications, network providers can take advantage of user context data for modeling user behavior in order to manage their resources and service allocations more effectively. Environmental monitoring applications are another type of emerging applications on mobile phones. Here, each mobile phone acts as a sensor for monitoring particular parameters of the user environment, for example, noise level

[Santini et al. 2009], CO2 footprints [Mun et al. 2009], and traffic monitoring [Mohan et al. 2008], which is then aggregated with data from other users to cover a larger area. As new sensors become available on mobile phones, more environmental metrics are envisioned to be monitored with these platforms. Currently, healthcare applications can be easily extended from personal monitoring to large-scale monitoring for epidemiological purposes. Recent advances in social signal processing (SSP) have paved the way for a new class of socially intelligent applications. The potential of what can be achieved by combining these techniques with mobile-phone-centric observations have been highlighted in a variety of recent studies (e.g., [Zhang et al. 2008; Eagle and Pentland 2006; Onnela et al. 2007]). Pioneers in the SSP field, such as Alex Pentland and Nathan Eagle, have emphasised that the "very nature of the mobile phone makes them an ideal vehicle to study both individuals and organizations" [Eagle and Pentland 2006]. Applications can take advantage of data captured by mobile phone-centric sensing for analyzing a spectrum of social networks ranging from personal and small groups to large-scale communities. The pervasive data entailing user behavior that can be gathered through such opportunistic sensing applications (e.g., reality mining [Eagle and Pentland 2006]) is an invaluable resource for human studies applications. It is likely in the near future that the use of mobile phones with pervasive sensing and social signal processing capabilities will share the current multimillion pound market for social surveys. Examples range from smaller scale studies such as organizational behavior [Cross et al. 2002]. to large-scale ones, such as the International Social Survey Programme [GESIS 2009] and the European Social Survey [European Social Survey 2009].

6.6. Conclusion

Recent advances in computing, storage, and wireless technology, together with the introduction of MEMS sensors, have made mobile phones ideal candidate platforms for ubiquitous observation of user context.

While today's smartphones have become increasingly multipurpose platforms, it is still a challenging task to add opportunistic sensing and context processing capabilities without jeopardizing the user's overall mobile phone experience.

A large diversity of different opportunistic sensing techniques and applications utilizing those have been developed in recent years. This article analyzes these approaches and represents a first attempt to classify the techniques and methodologies used as distinct components of a mobile sensing architecture. The resultant architecture comprises three major stages, namely sensing, preprocessing, and context recognition, involving a variety of techniques to fulfill one or more tasks inside each of the stages. For each of the stages, a thorough analysis of advantages and shortcomings of the currently implemented techniques has been provided. Ensuring adequate quality of context while considering device constraints for user context-recognition requires a deeper understanding of these and how they interact with each other. This article has contributed towards this understanding by deriving recommendations from published analyses of how these techniques at different stages of the architecture should be combined to build more reliable mobile sensing platforms.

The article concludes by highlighting remaining challenges for mobile-phone-based opportunistic sensing system. Examples are the development of techniques for resource- and context-aware sensing and feature selection, labeling of complex and unknown events, and the preservation of user privacy during sensing process. Developing concrete solutions for these issues will open the doors to countless novel applications exploiting those capabilities, making opportunistic mobile sensing systems key elements of future service environments.

APPENDIX

Table II. Comparison of several systems prototypes for user context recognition

System	Sensor(s)	Sensor Node	Processing Units	Sampling Rate (Hz)	Preprocessing	Context Inference	Context	Accuracy	Goal
SurroundSense [Azizyan et al. 2009]	Camera, Microphone, Accelerometer, WiFi	Nokia N95	Nokia N95	Accelerometer (<4), Camera (.2), Microphone 8Khz, WiFi (.2)	Normalization, Average, Variance, HSL, Color, Light, Noise level	SVM, K-means clustering, Thresholds	Environment, User motion	87%	Localization via ambient fingerprints
UPCASE [Santos et al. 2009]	Tri-axial accelerometer, Humidity, Light, Temperature, Microphone, GPS	Blue Sentry module	Smartphone (Nokia N95, Sony Ericson W910i)	(<20, <4) Accelerometers	Variance FFT, Thresholds mean	Decision tree C4.5 ID3	Walking, Running, Standing, Lying, Being inside or out side	C4.5>90 %, ID3>91%	Recognizing user context
CenceMe [Miluzzo et al. 2008]	Microphone, Accelerometer, Bluetooth, GPS, Camera	Nokia N95	Smartphone + backend server	Using poweraware duty cycle, audio, and accelerometer (.1 to .01) GPS and Bluetooth (.01 to .001)	DFT, Mean, Standard deviation, Number of peaks	Decision tree J48, K-means clustering, (Smartphone, on) thresholds, JRIP rule learning (on backend server)	Walking, Running, Standing, Presence of conversation, Mobile phones in vicinity, Mobility, Social context	Classification of different features varies with different position of the phone and environment	Detect user social presence to publish on social networking applications
[Kwapisz et al. 2011]	Accelerometer	Smartphone	Backend server	20	Average standard deviation, Average absolute difference, Average resultant acceleration, Time between peaks, Binned distribution	J48, Logistic Regression,	Walking, Jogging, Upstairs, Downstairs, Sitting, Standing	Walking and jogging >90% generally	Activity recognition using mobile phone embedded accelerometer
[Lester et al. 2006]	Microphone, Compass, Accelerometer, Temperature, Humidity sensor, and etc.	Multimodal sensor board (MSB)	Backend server	4	Cepstral coefficients, Log FFT frequency bands, Spectral entropy, Energy, Mean, Linear FFT Frequency bands, Correlation coeffs, Integration	HMM	Walking down stairs, Sitting, Riding elevator down, Riding elevator up, Brushing teeth	90%	Providing genetic, personal activity recognition system

Table II. Continue

System	Sensor(s)	Sensor Node	Processing Units	Sampling Rate (Hz)	Preprocessing	Context Inference	Context	Accuracy	Goal
SenSay [Siewiorek et al. 2003]	Microphones, GPS, 2-axis accelerometer, Blue- Spoon headset, Internal clock	Sensor box as central hub and wearable sensors	Notebook		Average, SAD, FFT, Normalization, Principal component analysis	Hierarchy of thresholds	User states as Idle, Uninter-ruptible, Active and default		Provides a contextaware mobile phone with dynamic adaptation to environment
Reality Mining [Eagle and Pentland 2006]	Bluetooth (BTID), GSM (cell towers ID)	Nokia 6600	Smartphone + backend server	Once every 5 min	Distribution (PDF), Entropy	HMM, Bayes rule, GMM	Location pattern, Proximity pattern	(95 %) Identify next location, 90% (face to face contacts), 90% (rela-tionships)	Social pattern in daily activity, Infer relationship, Human landmarks, Model organizational rhythm
Serendipity [Eagle and Pentland 2005]	Bluetooth (BTID), GSM (cell towers ID)	Nokia 3650	Smartphone + backend server	Once every 5 min	Updating thresholds and weights sent by user over backend server	GMM, Thresholds	Social location pattern, Social relation, Proximity, Similarity in profiles	Classification of different features varies with different position of the phone and environment	Detect user social networks of relationship, Cueing informal face-to-face interactions
[Anderson and Muller 2006a]	GSM receiver	Mobile phone (SPV c500)	SmartPhone (SPV C500)	.06	Mean, Variance	HMM, K-means	Walking, Stationary, Driving	80%	Context awareness by GSM signals
[Sohn et al. 2005]	GSM receiver	Mobile phone (Audiovox SMT 5600)	Backend server	1	Euclidean distance, Correlation coefficient, Number of common cells between two measurements, Mean, Variance	Boosted logistic regression	Walking, Running, Driving	85%	Recognizing high-level activities with coarse-grained GSM data

Table II. Continue

System	Sensor(s)	Sensor Node	Processing Units	Sampling Rate (Hz)	Preprocessing	Context Inference	Context	Accuracy	Goal
[Gyorbíró et al. 2009]	Accelerometer, Magnetometer, Gyroscope	Motion band	Smartphone (Nokia 6630)	50	Intensity, Normalization	Neural networks	Sitting, Typing, Gesticulating, Walking, Running, Cycling	79.76%	Recognizing motional activities via mobile phone
[Yang 2009]	Accelerometer	Smartphone (Nokia N95)	Mobile phone/PC	36, 0.1	Moving average filtering, Mean, Standard deviation from horizontal and vertical axis (for mobile use)	Decision tree (C4.5), K-means clustering, HMM	Standing, Running, Walking, Biking, Driving, Sitting	66% With simplified features	Detecting physical activity with mobile phone to provide physical activity diary
[Kawahara et al. 2007]	Accelerometer	Mobile phone	Backend server	20	Variance, Average, FFT, Sensor angle	Thresholds	Physical activities: Sitting, Standing, Running, and Leaning. Phone position: Chest pocket, Trousers pocket, and Not taken by user	96% >	Detecting user activity with mobile handset
InSense [Blum et al. 2006]	Accelerometer, Microphone, Camera, WiFi	External sensors	PDA (Sharp Zaurus SL6000L)	Accelerometer (90), Microphone (8), WiFi (.01), Camera (.16)	Mean, Variance, Spectral entropy, Energy maximum and number of autocorrelation peaks	Naïve Bayes classifier using Gaussian probability distribution	Location, Activity, Posture, Speech	>73%	Real-time context recognition and user interest prediction
MobSens [Kanjo et al. 2009]	Air pollution sensor, Microphone, GSM, GPS	Smartphone (Nokia N95, N80) and External sensors	Smartphone		Filtering, Mapping		Pollution, Noise, Common location		Enabling environmental data collection from mobile phone

Table II. Continue

System	Sensor(s)	Sensor Node	Processing Units	Sampling Rate (Hz)	Preprocessing	Context Inference	Context	Accuracy	Goal
EEMSS [Wang et al. 2009]	Accelerometer, Microphone, GPS	Nokia N95	Nokia N95	0.1s (Accelerometer), 0.5–10sec (Microphone)	Standard Deviation, FFT	Decision tree	Walking, Vehicle, Resting, Home talking, Home entertaining, Working, Meeting, Office_ loud, Place_ quiet, Place _speech and Place_ loud	92.56% with a standard deviation of 2.53%	Providing an energyefficient sensing system for mobile phones
AniDiary [Cho et al. 2007]	GPS, Phone usage	Smartphone	Smartphone/ PC	Average, Maximum, Minimum, and Frequency	Bayesian networks	Context as Place-activity, Emotional/ conditional, Circumstantial/ situational, Events.	75%	To represent user daily life with cartoonbased information collected via mobile devices like Smartphone	
Soundsense [Lu et al. 2009]	Microphone	Apple Iphone	Smartphone	8,000	Zero crossing rate, Low energy frame rate, Spectral rolloff, Spectral centroid, Bandwidth, Normalized weighted phase deviation, Relative spectral entropy and Mel frequency ceptral coefficient, Spectral variance	Markov model, Decision tree (J48), Gaussian discriminative model	Human voice, Music, Ambient	>78%	Recognizing everyday life sound events on mobile phone

ACKNOWLEDGMENTS

The authors would like to thank Dr. Lloyd Wood for his kind support in the final proof reading of this article. Your comments, editorial updates, and discussions were invaluable to us. In particular, we appreciate the time you dedicated, despite the few remaining days that you had left with us. The authors would also like to thank the reviewers for their insightful comments and suggestions on the earlier drafts of this article.

REFERENCES

ALLEN, F. R., AMBIKAIRAJAH, E., LOVELL, N. H., AND CELLER, B. G. 2006. An adapted Gaussian mixture model approach to accelerometry-based movement classification using time-domain features. In *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. (EMBS'06)*, 3600–3603.

AMINIAN, K., ROBERT, P., JÉQUIER, E., AND SCHUTZ, Y. 1995. Incline, speed, and distance assessment during unconstrained walking. *Med. Sci. Sports Exercise 27*, 2, 226–34.

ANDERSON, I. AND MULLER, H. 2006a. Practical context awareness for GSM cell phones. In *Proceedings of the 10th IEEE International Symposium on Wearable Computers.* 127–128.

ANDERSON, I. AND MULLER, H. 2006b. Practical activity recognition using GSM data. Tech. rep. CSTR-06-016 Department of Computer Science, University of Bristol.

ANDERSON, I., MAITLAND, J., SHERWOOD, S., BARKHUUS, L., CHALMERS, M., HALL, M., BROWN, B., AND MULLER, H. 2007. Shakra: Tracking and sharing daily activity levels with unaugmented mobile phones. *Mob. Netw. Appl. 12*, 185–199.

ANTONSSON, E. K. AND MANN, R. W. 1985. The frequency content of gait. *J. Biomech. 18*, 39–47.

AOKI, P. M., GRINTER, R. E., HURST, A., SZYMANSKI, M. H., THORNTON, J. D., AND WOODRUFF, A. 2002. Sotto voce: Exploring the interplay of conversation and mobile audio spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Changing our world, changing ourselves*. 431–438.

ARIKAWA, M., KONOMI, S., AND OHNISHI, K. 2007. Navitime: Supporting pedestrian navigation in the real world. *IEEE Perv. Comput. 6*, 21–29.

ASAHI KASCI. 2006. 6 Axis Electronic Compass Chip - AK8976A - Asahi Kasei. 2006. `http://embedded-system.net/6-axis-electronic-compass-chip-ak8976a-asahi-kasei.html`. (Last accessed 7/07).

ATALLAH, L., LO, B., ALI, R., KING, R., AND YANG, G., 2009. Real-time activity classification using ambient and wearable sensors. *IEEE Tran. Inform. Technol. Biomed. 13*, 6, 1031–1039.

ATALLAH, L. AND YANG, G. 2009. The use of pervasive sensing for behavior profiling – a survey. *Perv. Mobile Comput. 5*, 5, 447–464.

AZIZYAN, M., CONSTANDACHE, I., AND ROY CHOUDHURY, R. 2009. SurroundSense: Mobile phone localization via ambience fingerprinting. In *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking*. 261–272.

BAHL, P. AND PADMANABHAN, V. N. 2000. RADAR: An in-building RF-based user location and tracking system. In *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies*. Vol. 2, 775–784.

BAO, L. AND INTILLE, S. 2004. Activity recognition from user-annotated acceleration data. In *Pervasive Computing*, A. Ferscha and F. Mattern, Eds., Lecture Notes in Computer Science, vol. 3001, Springer, Springer Berlin/Heidelberg, 1–17.

BARBEAU, S. J., WINTERS, P. L., GEORGGI, N. L., LABRADOR, M. A., AND PEREZ, R. 2010. Travel assistance device: Utilising global positioning system-enabled mobile phones to aid transit riders with special needs. *Intell. Transport Syst. IET 4*, 12–23.

BAR-NOY, A. AND KESSLER, I. 1993. Tracking mobile users in wireless communications networks. In *Proceedings of the 12th Annual Joint Conference of the IEEE Computer and Communications Societies. Networking: Foundation for the Future*. Vol. 3, 1232–1239.

BASU, S. 2003. A linked-HMM model for robust voicing and speech detection. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*. I-816–I-819.

BERCHTOLD, M., BUDDE, M., GORDON, D., SCHMIDTKE, H. R., AND BEIGL, M. 2010. ActiServ: Activity recognition service for mobile phones. In *Proceedings of the International Symposium on Wearable Computers (ISWC)*. 1–8.

BHATTACHARYA, A., MCCUTCHEON, E. P., SHVARTZ, E., AND GREENLEAF, J. E. 1980. Body acceleration distribution and O2 uptake in humans during running and jumping. *J. Appl. Physiol. 49*, 881–887.

BHATTACHARYA, A. AND DAS, S. K. 1999. LeZi-update: An information-theoretic approach to track mobile users in PCS networks. In *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking*. 1–12.

BIEBER, G., VOSKAMP, J., AND URBAN, B. 2009. Activity recognition for everyday life on mobile phones. In *Proceedings of the Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, C. Slephanidis, Ed., Lecture Notes in Computer Science, vol. 5615, Springer, Springer Berlin/Heidelberg, 289–296.

BLANKE, U. AND SCHIELE, B. 2008. Sensing location in the pocket. In *Proceedings of the 10th International Conference on Ubiquitous Computing* (UbiComp, adjunct proceedings). 1–2.

BLUM, M., PENTLAND, A., AND TROSTER, G. 2006. InSense: Interest-based life logging. *IEEE Multimedia 13*, 40–48.

BOUTEN, C. V. C., KOEKKOEK, K. T. M., VERDUIN, M., KODDE, R., AND JANSSEN, J. D. 1997. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Trans. Biomed. Eng. 44,* 136–147.

BREZMES, T., GORRICHO, J., AND COTRINA, J. 2009. Activity recognition from accelerometer data on a mobile phone. In *Proceedings of the 10th International Work-Conference on Artificial Neural Networks: Part II: Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living*. 796–799.

BRUNS, E., BROMBACH, B., ZEIDLER, T., AND BIMBER, O. 2007. Enabling mobile phones to support large-scale museum guidance. *IEEE Multimedia 14*, 2, 16–25.

CAKMACI, O. AND COUTAZ, J. 2002. Context awareness in systems with limited resources. In *Proceedings of the 3rd Workshop on Artificial Intelligence in Mobile Systems (AIMS).*

CANDIA, J., GONZÁLEZ, M. C., WANG, P., SCHOENHARL, T., MADEY, G., AND BARABÁSI, A.-L. 2008. Uncovering individual and collective human dynamics. *J. Physics A, 41*, 22, 4015–26.

CAPPÉ, O., MOULINES, E., AND RYDÉN, T. 2005. *Inference in Hidden Markov Models*. Springer, Berlin.

CAPPOZZO, A. 1989. Low frequency self-generated vibration during ambulation in normal men. *J. Biomech. 15,* 599–609.

CARUANA, R. AND NICULESCU-MIZIL, A. 2006. An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd International Conference on Machine Learning.* 161–168.

CHEN, D. M., TSAI, S. S., VEDANTHAM, R., GRZESZCZUK, R., AND GIROD, B. 2009. Streaming mobile augmented reality on mobile phones. In *Proceedings of the 8th IEEE International Symposium on Mixed and Augmented Reality.* 181–182.

CHENG, Y., CHAWATHE, Y., LAMARCA, A., AND KRUMM, J. 2005. Accuracy characterization for metropolitan-scale Wi-Fi localization. In *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services.* 233–245.

CHO, S. B., KIM, K. J., HWANG, K. S., AND SONG, I. J. 2007. AniDiary: Daily cartoon-style diary exploits Bayesian networks. *J. Perv. Comput. 6*, 3, 67–74.

CHO, S. J., OH, J. K., AND BANG, W. C. 2004. Magic wand: A hand-drawn gesture input device in 3-D space with inertial sensors. In *Proceedings of the 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR'9)*. 106–11.

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. A., LESTER, J., WYATT, D., AND HAEHNEL, D. 2008. The mobile sensing platform: An embedded activity recognition system. *IEEE Perv. Comput.* 7, 32–41.

CRK, I., ALBINALI, F., GNIADY, C., AND HARTMAN, J. 2009. Understanding energy consumption of sensor enabled applications on mobile phones. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. (EMBC'09)*. 6885–6888.

CROSS, R., BORGATTI, S., AND PARKER, A. 2002. Making invisible work visible : Using social network analysis to support strategic collaboration. *Calif. Manage. Rev. 44*, 2, 25–46.

DARTMOUTH COLLEGE. 2010. Mobile sensing group. `http://sensorlab.cs.dartmouth.edu./`. (Last accessed 10/10).

DAS, T., MOHAN, P., PADMANABHAN, V. N., RAMJEE, R., AND SHARMA, A. 2010. PRISM: Platform for remote sensing using smartphones. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and services.* 63–76.

DAVIS, M., SMITH, M., STENTIFORD, F. W. M., BAMIDELE, A., CANNY, J., GOOD, N., KING, S., AND JANAKIRAMAN, R. 2006. Using context and similarity for face and location identification. In *Proceedings of the IS & T/SPIE Electronic Imaging Conference.* 1–9.

DELIGNE, S., DHARANIPRAGADA, S., GOPINATH, R., MAISON, B., OLSEN, P., AND PRINTZ, H. 2002. A robust high accuracy speech recognition system for mobile applications. *IEEE Trans. Speech Audio Process. 10*, 8, 551–561.

DESELAERS, T., HEIGOLD, G., AND NEY, H. 2008. SVMs, Gaussian mixtures, and their generative/discriminative fusion. In *Proceedings of the 19th International Conference on Pattern Recognition (ICPR'08)*. 1–4.

DIXON, S. 2006. Onset detection revisited. In *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx06).* 18–20.

DUDA, R. O., HART, P. E., AND STORK, D. G. 2000. *Pattern Classification* 2nd Ed. John Wiley and Sons Hoboken, NJ.

EAGLE, N. AND PENTLAND, A. 2005. Social serendipity: Mobilizing social software. *IEEE Perv. Compu. 4,* 28–34.

EAGLE, N. AND PENTLAND, A. 2006. Reality mining: Sensing complex social systems. *Pers. Ubiq. Comput. 10*, 4, 255–268.

EAGLE, N., QUINN, J. A., AND CLAUSET, A. 2009. Methodologies for continuous cellular tower data analysis. In *Proceedings of the 7th International Conference on Pervasive Computing.* Lecture Notes in Computer Science, vol. 5538, Springer-Verlag, Berlin, Heidelberg, 342–353.

ERMES, M., PARKKA, J., AND CLUITMANS, L. 2008. Advancing from offline to online activity recognition with wearable sensors. In *Proceedings of the 30th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society.* 4451–4454.

EROL, B., GRAHAM, J., ANTÚNEZ, E., AND HULL, J. J. 2008. HOTPAPER: Multimedia interaction with paper using mobile phones. In *Proceeding of the 16th ACM International Conference on Multimedia.* 399–408.

EUROPEAN SOCIAL SURVEY. 2009. http://www.europeansocialsurvey.org./. (Last accessed 10/11).

EVANS, A., DUNCAN, G., AND GILCHRIST, W. 1991. Recording accelerations in body movements. *Med. Bio. Eng. Comput. 29*, 1, 102–104.

FANG, L., ANTSAKLIS, P. J., MONTESTRUQUE, L. A., MCMICKELL, M. B., LEMMON, M., SUN Y., FANG, H., KOUTROULIS, I., HAENGGI, M., XIE, M., AND XIE, X. 2005. Design of a wireless assisted pedestrian dead reckoning system - the NavMote experience. *IEEE Trans. Instrument. Measure. 54*, 6, 2342–2358.

FAWCETT, T., AND PROVOST, F. 1996. Combining data mining and machine learning for effective user profiling. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD96).* 8–13.

FERRO, E. AND POTORTI, F. 2005. Bluetooth and Wi-Fi wireless protocols: A survey and a comparison. *IEEE Wireless Communi. 12,* 12–26.

FLEURY, A., VACHER, M. AND NOURY, N. 2010. SVM-based multimodal classification of activities of daily living in health smart homes: Sensors, algorithms, and first experimental results. *IEEE Trans. Inform. Techno. Biomed. 14*, 2, 274–283.

FRIEDMAN, J. H. 1997. On bias, variance, 0/1–loss, and the curse-of-dimensionality. *Data Mining Knowl. Discovery 1,* 55–77.

FRIGO, M. 1999. A fast Fourier transform compiler. In *Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation.* 169–180.

FROEHLICH, J., CHEN, M. Y., CONSOLVO, S., HARRISON, B., AND LANDAY, J. A. 2007. My Experience: A system for in situ tracing and capturing of user feedback on mobile phones. In *Proceedings of the 5th International Conference on Mobile Systems, Applications and Services.* 57–70.

GAONKAR, S., LI, J., CHOUDHURY, R. R., COX, L., AND SCHMIDT, A. 2008. Micro-Blog: Sharing and querying content through mobile phones and social participation. In *Proceeding of the 6th International Conference on Mobile Systems, Applications, and Services.* 174–186.

GESIS. 2009. Leibniz Institute for the Social Sciences. http://www.gesis.org/en/services/data/survey-data/issp/modules-study-overview/social-networks/2001/. (Last accessed 10/11).

GHIANI, G. AND PATERNÒ, F. 2010. Supporting mobile users in selecting target devices. *J. Univ. Compu. Sci. 16*, 15, 2019–2037.

GOLDING, A. AND LESH, N. 1999. Indoor navigation using a diverse set of cheap, wearable sensors. In *Procceedings of the International Symposium on Wearable Computers (ISWC'99).* 29–36.

GONZALEZ, M. C., HIDALGO, C. A., AND BARABASI, A. 2008. Understanding individual human mobility patterns. *Nature 453,* 779–782.

GRISWORLD, W. G., BOYER, R., BROWN, S. W., TRUONG, T. M., BHASKER, E., JAY, G. R., AND SHAPIRO, R. B. 2002. Using mobile technology to create opportunitistic interactions on university campus. In *Proceedings of the Ubicomp Workshop on Supporting Spontaneous Interaction in Ubiquitous Computing Settings*.

GU, J., MUKUNDAN, R., AND BILLINGHURST, M. 2008. Developing mobile phone AR applications using J2ME. In *Proceedings of the 23rd International Conference Image and Vision Computing.* 1–6.

GYORBÍRÓ, N., FÁBIÁN, A., AND HOMÁNYI, G. 2009. An activity recognition system for mobile phones. *Mobile Netw. Appl. 14,* 82–91.

HAILES, S., SICARI, S., AND ROUSSOS, G., EDS. 2009. *Sensor Systems and Softwear* 1st Ed. Springer, New York, NY.

Häkkilä, J., and Chatfield, C. 2005. 'It's like if you opened someone else's letter': user perceived privacy and social practices with SMS communication. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services.* 219–222.

Haro, A., Mori, K., Setlur, V., and Capin, T. 2005. Mobile camera-based adaptive viewing. In *Proceedings of the 4th International Conference on Mobile and Ubiquitous Multimedia.* 78–83.

Havasi, C., Speer, R., Pustejovsky, J., and Lieberman, H. 2009. Digital intuition: Applying common sense using dimensionality reduction. *IEEE Intell. Syst. 24*, 24–35.

Haykin, S. S. 2009. *Neural Networks and Learning Machines*. Prentice Hall, Upper Saddle River, NJ.

He, J., Li, H., and Tan, J. 2007. Real-time daily activity classification with wireless sensor networks using hidden markov model. In *Proceedings of the 29th Annual International Conference on Engineering in Medicine and Biology Society (EMBS'07)*. 3192–95.

Herrera, J. C., Work, D. B., Herring, R., Ban, X. J., and Bayen, A. M. 2010. Evaluation of traffic data obtained via GPS-enabled mobile phones: The mobile century field experiment. *Transport. Res. Part C: 18*, 4, 568–83.

Hightower, J. and Borriello, G. 2001. Location systems for ubiquitous computing. *IEEE Comput. Maga. 4*, 8, 57–66.

Himberg, J., Korpiaho, K., Mannila, H., Tikanmaki, J., and Toivonen, H. T. T. 2001. Time series segmentation for context recognition in mobile devices. In *Proceedings of the IEEE International Conference on Data Mining, (ICDM'01)*. 203–210.

Hoseinitabatabaei, S. A., Gluhak, A., and Tafazolli, R. 2011. uDirect: A novel approach for pervasive observation of user direction with mobile phones. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 74–83.

Huynh, T., Fritz, M., and Schiele, B. 2008. Discovery of activity patterns using topic models. In *Proceedings of the 10th International Conference on Ubiquitous Computing.* 10–19.

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

Iso, T. and 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.* 141–144.

Jain, A. K., Murty, M. N. and Flynn, P. J. 1999. Data clustering: A review. *ACM Comput. Sur. 31*, 264–323.

Kanjo, E., Bacon, J., Roberts, D., and Landshoff, P. 2009. MobSens: Making smart phones smarter. *IEEE Perv. Comput. 8,* 50–57.

Kanjo, E. 2010. NoiseSPY: A real-time mobile phone platform for urban noise monitoring and mapping. *Mob. Netw. Appl. 15,* 562–574.

Kansal, A., and Zhao, F. 2007. Location and mobility in a sensor network of mobile phones. In *Proceedings of the 17th International Workshop on Network and Operating Systems Support for Digital Audio & Video (NOSSDAV)*.

Kapadia, A., Kotz, D., and Triandopoulos, N. 2009. Opportunistic sensing: Security challenges for the new paradigm. In *Proccedings of the International Communication Systems and Networks and Workshops. (COMSNETS'09)*, 1–10.

Katzakis, N. and Hori, M. 2009. Mobile phones as 3-DOF controllers: A comparative study. In *Proceedings of the 8th IEEE International Conference on Dependable, Autonomic and Secure Computing (DASC'09)*. 345–349.

Kawahara, Y., Kurasawa, H., and Morikawa, H. 2007. Recognizing user context using mobile handsets with acceleration sensors. In *Proceedings of the IEEE International Conference on Portable Information Devices (PORTABLE'07)*. 1–5.

Kim, S. J. and Lee, C. Y. 1996. Modeling and analysis of the dynamic location registration and paging in microcellular systems. *IEEE Trans. Vehicular Technol. 45*, 82–90.

Kim, T., Chang, A., Holland, L., and Pentland, A. 2008. Meeting mediator: Enhancing group collaboration with sociometric feedback. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW'08)*. 457–466.

Könönen, V., Mäntyjärvi, J., Similä, H., Pärkkä, J., and Ermes, M. 2010. Automatic feature selection for context recognition in mobile devices. *Perv. Mobile Compu. 6*, 181–197.

Kourogi, M. and Kurata, T. 2003a. Personal positioning based on walking locomotion analysis with self-contained sensors and a wearable camera. In *Proceedings of the 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality.* 103–112.

KOUROGI, M. AND KURATA, T. 2003b. A wearable augmented reality system with personal positioning based on walking locomotion analysis. In *Proceedings of the 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality.* 342–343.

KRUMM, J. AND HORVITZ, E. 2004. LOCADIO: Inferring motion and location from Wi-Fi signal strengths. In *Proceedings of the Ist Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'04)*. 4–13.

KUNZE, K. AND LUKOWICZ, P. 2007. Using acceleration signatures from everyday activities for on-body device location. In *Proceedings of the 11th IEEE International Symposium on Wearable Computers.* 115–116.

KUNZE, K., LUKOWICZ, P., PARTRIDGE, K., AND BEGOLE, B. 2009. Which way am I facing: Inferring horizontal device orientation from an accelerometer signal. In *Proceedings of the International Symposium on Wearable Computers (ISWC'09)*. 149–150.

KUNZE, K., LUKOWICZ, P., JUNKER, H., AND TRÖSTER, G. 2005. Where am I: Recognizing on-body positions of wearable sensors. In *Proceedings of the Ist International Conference on Location- and Context-Awareness*, T. Strung and C. Linnhoyp-Popien, Eds., 257–268.

KWAPISZ, J. R., WEISS, G. M., AND MOORE, S. A. 2011. Activity recognition using cell phone accelerometers. *SIGKDD Explor. Newsl. 12*, 74–82.

KYRIAZAKOS, S. AND KARETSOS, G. 2000. Architectures for the provision of position location services in cellular networking environments. In *Proceedings of the Telecommunications and IT Convergence Towards Service E-volution*, J. Delgado, G. Stamoulis, A. Mulleey, D. Prevedourou, and K. Start, Eds., Lecture Notes in Computer Science, vol. 1774, Springer, Berlin/Heidelberg, 269–282.

LAASONEN, K., RAENTO, M., AND TOIVONEN, H. 2004. Adaptive on-device location recognition. In *Proceedings of the 2nd International Conference on Pervasive Computing.* 287–304.

LAFORTUNE, M. A. 1991. Three dimensional acceleration of tibia during walking and running. *J. Biomech. 24*, 877–86.

LANE, N. D., MILUZZO, E., HONG LU, PEEBLES, D., CHOUDHURY, T., AND CAMPBELL, A. T. 2010. A survey of mobile phone sensing. *IEEE Commun. Mag. 48*, 140–150.

LARSEN, J. E., AND LUNIEWSKI, M. 2009. Using mobile phone contextual information to facilitate managing image collections. In *Proceedings of the Personal Information Management Workshop(PIM'09)*. 73–75.

LEE, S. W. AND MASE, K. 2001. Incremental motion-based location recognition. In *Proceedings of the 5th IEEE International Symposium on Wearable Computers.* 123–130

LEE, S. W. AND MASE, K. 2002. Activity and location recognition using wearable sensors. *Proceedings of the IEEE Perv. Comput. 1*, 24–32.

LERCH, A. 2009. Software-based extraction of objective parameters from music performances. Ph.D. Dissertation. Berlin Technical University.

LESTER, J., CHOUDHURY, T., AND BORRIELLO, G. 2006. A practical approach to recognizing physical activities. In *Proceedings of the 4th International Conference on Pervasive Computing*, K. Fishan, B. Schiele, P. Nixon, and A. Quigley, Eds., Springer Berlin/Heidelberg, 1–16.

LIM, J. H., LI, Y., YOU, Y., AND CHEVALLET, J.-P. 2007. Scene recognition with camera phones for tourist information access. In *Proceedings of the IEEE International Conference on Multimedia and Expo.* 100–103.

LIU, Q., MCEVOY, P., AND LAI, C. 2006. Mobile camera supported document redirection. In *Proceedings of the 14th Annual ACM International Conference on Multimedia.* 791–792.

LOTTE, F., CONGEDO, M., LÉCUYER, A., LAMARCHE, F., AND ARNALDI, B. 2007. A review of classification algorithms for EEG-based brain-computer interfaces. *J. Neural Eng, 4*, 2, 1–13.

LU, H., PAN, W., LANE, N. D., CHOUDHURY, T., AND CAMPBELL, A. T. 2009. SoundSense: Scalable sound sensing for people-centric applications on mobile phones. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services.* 165–178.

LU, H., YANG, J., LIU, Z., LANE, N. D., CHOUDHURY, T. AND 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.* 71–84.

LUINGE, H. J., VELTINK, P. H., AND BATEN, C. T. M. 1999. Estimation of orientation with gyroscopes and accelerometers. In *Proceedings of the 1st Joint BMES/EMBS Conference Saving Humanity, Advancing Technology*, vol. 2, 844.

MARTENS, H. AND NAES, T. 2002. *Multivariate Calibration*. John Wiley & Sons, Hoboken, NJ.

MARTIN, G. N., CARLSON, N. R., AND BUSKIST, W. 2007. *Psychology*. Pearson Education, Harlow, Essex.

MATHIE, M. J., COSTER, A. C., LOVELL, N. H., AND CELLER, B. G. 2004. Accelerometery: Providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiol. Measure. 25*, 2, 1–20.

MAURER, U., ROWE, A., SMAILAGIC, A., AND SIEWIOREK, D. 2006. Location and activity recognition using eWatch: A wearable sensor platform. In *Proceedings of the Ambient Intelligence in Everyday Life*, Y. Cay and J. Abascal, Eds., Lecture Notes in Computer Science, vol. 3864, Springer Berlin/Heidelberg, 86–102.

MAYAGOITIA, R. E., LÖTTERS, J. C., VELTINK, P. H., AND HERMENS, H. 2002. Standing balance evaluation using a triaxial accelerometer. *Gait & Posture 16*, 55–59.

MCNEILL, P. AND CHAPMAN, S. 2005. *Research Methods*. Routledge, New York, NY.

MEEUWISSEN, E., REINOLD, P., AND LIEM, C. 2007. Inferring and predicting context of mobile users. *Bell Labs Tech. J. 12,* 79–86.

MILUZZO, E., LANE, N. D., FODOR, K., PETERSON, R., LU, H., MUSOLESI, M., EISENMAN, S. B., ZHENG, X., AND 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.* 337–350.

MIRKIN, B. 2005. *Clustering for Data Mining, a Data Recovery Approach*. Chapman & Hall/CRC, Boca Raton, FL.

MISHRA, A. R. 2004. *Fundamentals of Cellular Network Planning and Optimization*. John Wiley & Sons, Chichester, West Sussex.

MIZELL, D. 2003. Using gravity to estimate accelerometer orientation. In *Proceedings of the 7th IEEE International Symposium on Wearable Computers.* 252–253.

MOHAN, P., PADMANABHAN, V. N., AND RAMJEE, R. 2008. Nericell: Using mobile smartphones for rich monitoring of road and traffic conditions. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems.* 357–358.

MORRIS, S. J. AND PARADISO, J. A. 2002. Shoe-integrated sensor system for wireless gait analysis and real-time feedback. In *Proceedings of the 2nd Joint IEEE EMBS and BMES Conferences*, vol. 3, 2468–2469.

MOSTEFAOUI, G. K., PASQUIER-ROCHA, J., AND BREZILLON, P. 2004. Context-aware computing: A guide for the pervasive computing community. In *Proceedings of the IEEE/ACS International Conference on Pervasive Services (ICPS'04)*. 39–48.

MOTANI, M., SRINIVASAN, V., AND NUGGEHALLI, P. S. 2005. PeopleNet: Engineering a wireless virtual social network. In *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking.* 243–257.

MUN, M., REDDY, S., SHILTON, K., YAU, N., BURKE, J., ESTRIN, D., HANSEN, M., HOWARD, E., WEST, R., AND BODA, P. 2009. PEIR, the personal environmental impact report, as a platform for participatory sensing systems research. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services.* 55–68.

MURRAY, M. 1967. "Gait as a total pattern of movement". *Am. J. Phys. Med. 46*, 1, 290–333.

NOKIA 2005. *Proceedings of the Workshop on Large-Scale Sensor Networks and Their Applications.* 3–6.

OFSTAD, A., NICHOLAS, E., SZCODRONSKI, R., AND CHOUDHURY, R. R. 2008. AAMPL: Accelerometer augmented mobile phone localization. In *Proceedings of the Ist ACM International Workshop on Mobile Entity Localization and Tracking in GPS-Less Environments.* 13–18.

OLMEDILLA, D., FRÍAS-MARTÍNEZ, E., AND LARA, R. 2010. Mobile web profiling: A study of off-portal surfing habits of mobile users. In *User Modeling, Adeptation, and Personalization,* P. D. Bra, A. Kobra, and D. N. Chen, Eds., Lecture Notes in Computer Science, vol. 6075, Springer, 339–350.

OHMORI, N., NAKAZATO, M., AND HARATA, N. 2005. GPS mobile phone-based activity diary survey. In *Proceedings of the Eastern Asia Society for Transportation Studies.* 1104–1115.

OLGUIN, D. O. AND PENTLAND, A. 2006. Human activity recognition: Accuracy across common locations for wearable sensors. In *Proceedings of the IEEE 10th Symposium on Wearable Computers.*

OLGUIN, D. O. AND PENTLAND, A. 2008. Social sensors for automatic data collection. In *Proceedings of the 14th Americas Conference on Information Systems.* 1–10.

ONNELA, J. P., SARAMÄKI, J., HYVÖNEN, J., SZABO, G., LAZER, D., KASKI, K., KERTESZ, J., AND BARABASI, A. L. 2007. Structure and tie strengths in mobile communication networks. In *Proceedings of the National Academy of Science of the United States of America (PNAS).* 7332–36.

PAULOS, E. AND GOODMAN, E. 2004. The familiar stranger: Anxiety, comfort, and play in public places. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 223–230.

PEEBLES, D., LU, H., LANE, N. D., CHOUDHURY, T., AND CAMPBELL, A. T. 2010. Community-guided learning: Exploiting mobile sensor users to model human behavior. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence.*

PENTLAND, A. 2009. Honest Signals: How They Shape Our World. MIT Press, Cambridge, MA.

PICARD, R. W. AND HEALEY, J. 1997. Affective wearables. In *Proceedings of the 1st International Symposium on Wearable Computers.* 90–97.

PIETQUIN, O. 2004. A framework for unsupervised learning of dialogue strategies. P.D. dissertation. Universitaires de Louvain. ISBN: 2-930344-63-6.

PREECE, S. J., GOULERMAS, J. Y., KENNEY, L. P. J., HOWARD, D., MEIJER, K., AND CROMPTON, R. 2009. Activity identification using body-mounted sensors - a review of classification techniques. *Physiol. Measure. 30*, 4, 1–33.

LEVI, R. AND JUDD, T. 1996. Dead reckoning navigational system using accelerometer to measure foot impacts. U.S. patent 5583776, filed March 16, 1995, and issued December 10, 1996.

RAVI, N., SHANKAR, P., FRANKEL, A., ELGAMMAL, A., AND IFTODE, L. 2006. Indoor localization using camera phones. In *Proceedings of the 7th IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'06)*. 1–7.

RAVI, N., DANDEKAR, N., MYSORE, P., AND LITTMAN, M. L. 2005. Activity recognition from accelerometer data. In *Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence.* vol. 3, 1541–1546.

REDMOND, D. AND HEGGE, F. 1985. Observations on the design and specification of a wrist-worn human activity monitoring system. *Behav. Res. Methods 17*, 659–669.

ROSS, T. 2004. *Fuzzy Logic with Engineering Applications* 2nd Ed. John Wiley & Sons, Chichester, West Sussex.

RUF, B. AND DETYNIECKI, M. 2009. Identifying paintings in museum galleries using camera mobile phones. In *Proceedings of the Singaporean-French IPAL Symposium.* 125–34.

SANTINI, S., OSTERMAIER, B., AND ADELMANN, R. 2009. On the use of sensor nodes and mobile phones for the assessment of noise pollution levels in urban environments. In *Proceedings of the 6th International Conference on Networked Sensing Systems.* 31–38.

SANTOS, A. C., CARDOSO, J. M. P., FERREIRA, D. R., DINIZ, P. C., AND CHÁINHO, P. 2010. Providing user context for mobile and social networking applications. *Perv. Mobile Comput. 6*, 324–341.

SANTOS, A. C., TARRATACA, L., CARDOSO, J. M. P., FERREIRA, D. R., DINIZ, P. C., AND CHAINHO, P. 2009. Context inference for mobile applications in the UPCASE project. In *Proceedings of the MobileWireless Middleware, Operating Systems, and Applications*, J. Bonnin, C. Giannelliand, and T. Magedanz, Eds., Lecture Notes of the Institute for Computer Sciences, Social Information and Telecommunications Engineering, vol. 7, Springer Berlin Heidelberg, 352–365.

SASHIMA, A., INOUE, Y., IKEDA, T., YAMASHITA, T., AND KURUMATANI, K. 2008. CONSORTS-S: A mobile sensing platform for context-aware services. In *Proceedings of the International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP'08)*. 417–422.

SAUNDERS, J. 1996. Real-time discrimination of broadcast speech/music. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (CASSP'96)*, vol. 2, 993–996.

SCHEIRER, E. AND SLANEY, M. 1997. Construction and evaluation of a robust multifeature speech/music discriminator. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. (ICASSP'97)*. vol. 2, 1331–1334.

SCHILLER, J. 2003. *Mobile Communications*. Pearson, Harlow, Essex.

SEKINE, M., TAMURA, T., AKAY, M., FUJIMOTO, T., TOGAWA, T., AND FUKUI, Y. 2002. Discrimination of walking patterns using wavelet-based fractal analysis. *IEEE Trans. Neural Syst. Rehab. Eng. 10*, 188–196.

SEKINE, M., TAMURA, T., FUJIMOTO, T., AND FUKUI, Y. 2000. Classification of walking pattern using acceleration waveform in elderly people. In *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* vol. 2, 1356–1359.

SIEWIOREK, D., SMAILAGIC, A., FURUKAWA, J., KRAUSE, A., MORAVEJI, N., REIGER, K., SHAFFER, J., AND WONG, F. L. 2003. SenSay: A context-awaremobile phone. In *Proceedings of the 7th IEEE International Symposium on Wearable Computers.* 248–249.

SOHN, T., LI, K. A., LEE, G., SMITH, I., SCOTT, J., AND GRISWOLD, W. G. 2005. Place-its: A study of location-based reminders on mobile phones. In *Proceedings of the 5th International Conference on Ubiquitous Computing (Ubicomp'05).* 232–50.

SOHN, T., VARSHAVSKY, A., LAMARCA, A., CHEN, M., CHOUDHURY, T., SMITH, I., CONSOLVO, S., HIGHTOWER, J., GRISWOLD, W. G., AND DE LARA, E. 2006. Mobility detection using everyday GSM traces. In *Proceedings of the 8th International Conference on Ubiquitous Computing (UbiComp).* 212–24.

STMICROELECTRONICS. `http://www.st.com/stonline/products/literature/ds/12726/lis302dl.htm`. (Last accessed 6/10).

STIEFMEIER, T., ROGGEN, D., TROSTER, G., OGRIS, G., AND LUKOWICZ, P. 2008. Wearable activity tracking in car manufacturing. *IEEE Perv. Comput. 7*, 42–50.

SUN, M. AND HILL, J. O. 1993. A method for measuring mechanical work and work efficiency during human activities. *J. Biomech. 26*, 229–241.

SUH, Y., SHIN, C., AND WOO, W. 2009. A mobile phone guide: spatial, personal, and social experience for cultural heritage. *IEEE Trans. on Consumer Electron. 55*, 2356–2364.

TAKACS, G., CHANDRASEKHAR, V., GELFAND, N., XIONG, Y., CHEN, W., BISMPIGIANNIS, T., GRZESZCZUK, R., PULLI, K., AND GIROD, B. 2008. Outdoors augmented reality on mobile phone using loxel-based visual feature organization. In *Proceeding of the 1st ACM International Conference on Multimedia Information Retrieval.* 427–434.

TITTERTON, D. H. AND WESTON, J. L. 2002. *Strapdown Inertial Navigation Technology* 2nd Ed. Institution of Electrical Engineers, NewYork, NY.

VAHDATPOUR, A., AMINI, N., AND SARRAFZADEH, M. 2011. On-body device localization for health and medical monitoring applications. In *Proceeding of the IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 37–44.

VAHDATPOUR, A., AMINI, N., AND SARRAFZADEH, M. 2009. Toward unsupervised activity discovery using multidimensional motif detection in time series. In *Proceedings of the 21st International Conference on Artificial Intelligence.* 1261–1266.

VIEIRA, M. R., FRÍAS-MARTÍNEZ, E., BAKALOV, P., FRÍAS-MARTÍNEZ, V., AND TSOTRAS, V. J. 2010. Querying spatio-temporal patterns in mobile phone-call databases. In *Proceeding of the 11th International Conference on Mobile Data Management (MDM).* 239–248.

VINCIARELLI, A., PANTIC, M., AND BOURLAND, H. 2009. Social signal processing : Survey of an emerging domain. *Image Vision Comput. 27*, 12, 1743–59.

WAGNER, D., REITMAYR, G., MULLONI, A., DRUMMOND, T., AND SCHMALSTIEG, D. 2010. Real-time detection and tracking for augmented reality on mobile phones. *IEEE Trans. Visual. Comput. Graphics 16*, 355–368.

WANG, J., ZHAI, S., AND CANNY, J. 2006. Camera phone based motion sensing: Interaction techniques, applications and performance study. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology.* 101–110.

WANG, Y., LIN, J., ANNAVARAM, M., JACOBSON, Q. A., HONG, J., KRISHNAMACHARI, B., AND 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.* 179–192.

WEBB, A. 1999. *Statistical Pattern Recognition*. John Wiley & Sons, Chichester, West Sussex.

WITTKE, M., JÄNEN, U., DURASLAN, A., CAKAR, E., STEINBERG, M., AND BREHM, J. 2009. Activity recognition using optical sensors on mobile phones. In *Proceeding of the GI-Tagung Informatik,* S. Fischer, E. Maehle, and R. Reischuk, Eds. 2181–2194.

WOODMAN, O. J. 2007. An introduction to inertial navigation. Tech. report. University of Cambridge. ISSN: 1476-2986.

WU, J. K., DONG, L., AND XIAO, W. 2007. Real-time physical activity classification and tracking using wearble sensors. In *Proceeding of the 6th International Conference on Information, Communications & Signal Processing*. 1–6.

XSENS. 2011. Xsens MTi : Miniature AHRS - Attitude and Heading Sensor - Xsens. `http://www.xsens.com/en/general/mti`. (Last accessed 10/11).

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.* 1–10.

YE, N. 2004. *The Hand Book of Data Mining*. Lawrence Erlbaurm Associates, Mahwah, NJ.

YI, J. S., CHOI, Y. S., JACKO, J. A., AND SEARS, A. 2005. Context awareness via a single device-attached accelerometer during mobile computing. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services.* 303–306.

YIM, Y. 2003. The state of cellular probs. Berkeley. California Partners for Advanced Transit and Highways (PATH), Institute of Transportation Studies, ISSN:1055-1425.

YIM, Y. AND CAYFORD, R. 2001. Investigation of vehicles as probes using global positioning system and cellular phone tracking. California Partners for Advanced Transit and Highways (PATH), Institute of Transportation Studies, Berkeley. ISSN: 1055–1417.

ZHANG, Y.-Y., ZHANG, W.-G., ZHAO, X.-X., AND YMAN, H.-M. 2009. Study on electronic image stabilization system based on MEMS gyro. In *Proceeding of the International Conference on Electronic Computer Technology.* 641–643.

ZHAO, Y. 2000. Mobile phone location determination and its impact on intelligent transportation systems. *IEEE Trans. Intell. Transport. Syst. 1*, 55–64.

ZHANG, S., YUAN, C., AND ZHANG, Y. 2008. Handwritten character recognition using orientation quantization based on 3D accelerometer. In *Proceedings of the 5th Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services.* 54:1–54:6.

ZINNEN, A., BLANKE, U., AND SCHIELE, B. 2009a. An analysis of sensor-oriented vs. model-based activity recognition. In *Proceeding of the International Symposium on Wearable Computers (ISWC'09)*. 93–100.

ZINNEN, A., WOJEK, C., AND SCHIELE, B. 2009b. Multi activity recognition based on bodymodel-derived primitives. In *Proceedings of the 4th International Symposium on Location and Context Awareness.* 1–18.

ZULKEFLY, S. N. AND BAHARUDIN, R. 2009. Mobile phone use amongst students in a university in malaysia: Its correlates and relationship to psychological health. *Euro. J. Sci. Res. 37*, 2, 206–18.

Received January 2011; revised July, October 2011; accepted November 2011