A Survey on Smartphone Based Systems for Opportunistic User Context Recognition

SEYED AMIR HOSEINI-TABATABAEI, ALEXANDER GLUHAK AND RAHIM TAFAZOLLI

Center for Communication Systems Research, University of Surrey

The ever growing computation and storage capability of mobile phones has 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. Particularly, as nonintrusive autonomous sensing and context recognition are one of the most desirable characteristics of a personal sensing system; commendable 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*, *pre-processing* and *context recognition* phases are introduced. The main techniques used for the realization of the respective processes during these phases are described, highlighting strengths and limitations of those. 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

1. INTRODUCTION

Efforts to understand human behaviour date back to the early physiological, psychological and sociologic studies of the 18th and 19th centuries. Since then, different branches of science with different perspectives have studied human behaviour in terms of relations between different causes, events and types of behaviour. A brief look at scientific theories about humans shows that the causes of the behaviour involves 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 behaviour is rooted in the combination of these causes, a single perspective can never give a comprehensive explanation of behaviour. When we add to this fact the uniqueness of an individual, understanding human behaviour 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 behaviour based upon the correlation between what individuals expose to their environment and a specific type of behaviour. Observation of such correlations can thus be utilised to develop a model for human behaviour in various situations. Studies in human behaviour show that a person's behaviour is highly dependent on perception, context, environment, prior knowledge and interaction with others [Attalah and Yang, 2009]. In this regard, various studies (e.g. [Attalah and Yang, 2009], [Aoki et al., 2002], [Suh et al., 2009]) have concluded that in order to model human behaviour, 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 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, analysing behaviour of a group of participates during a study or healthcare approaches where patients would accept wearing the device for a long period of time. The collection of longer term user context information with ubiquitous coverage thus still remains an open technological challenge.

In the 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, mobile phone-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, modelling 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 .Hence techniques developed for one of these systems are applicable for 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 pre-processing 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, e.g. 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 of the 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. According 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 realising 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 organised 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 pre-processing, discussing recent advances and techniques for calibration and feature extraction. Section 4 investigates the currently widely implemented algorithms for the context recognition phase and analysis 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 modelling and understanding different facets of human behaviour. Examples range from analysis of the pattern of message communication, 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 advancement has taken place only recently. In their recent survey on mobile phone sensing [Lane et al., 2010], Lane et al. 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 the developers to take advantage of resources in back-end servers as well as for analysing and collecting data from a large number of users. 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 realise 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 pre-processing step. The refined data or features extracted through pre-processing 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 battery powered portable handsets, make the configuration of sensing very challenging. The main challenge that impacts the sensing stage is to accurately recognise the required context with a minimum number of sensors and sensing frequency. Through pre-processing, 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 pre-processing and classification techniques to less computational intensive methods. During the pre-processing stage, redundancy and noise are minimised 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 performed often offline by frequent training and evaluation of classification performance, with the aim to improve the classification accuracy by discarding indiscriminative or highly correlated features and to 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 architecture require a feature extraction control mechanisms to exploit these tradeoffs 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 less privacy concerns and reduces the power required for transferring the data. However, the complexity of the applications is limited by the local computing and storage resources. Uploading data to a back end server meanwhile provides better opportunities for the exploitation of aggregate data from a large number of users, while allowing for the realisation of more complex applications. However, it requires more serious considerations for privacy and power consumption for the remote context delivery.

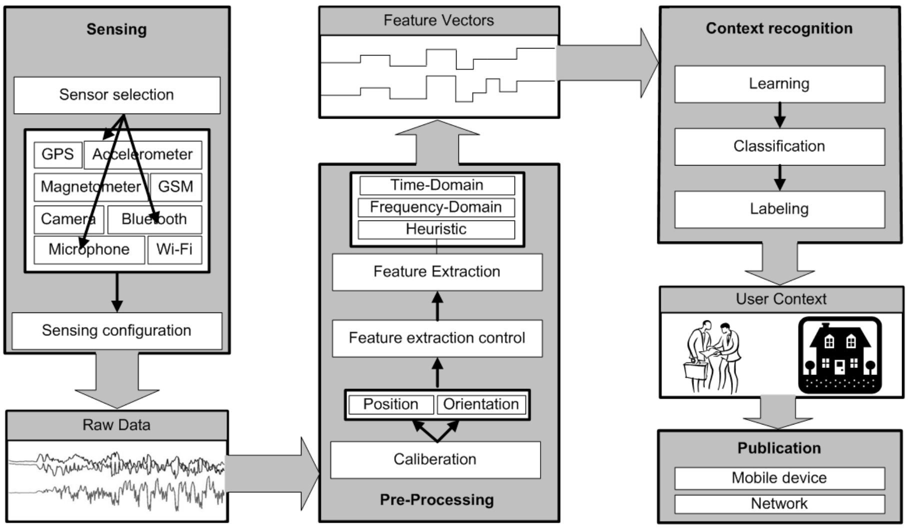


Figure 1. Overview of tasks and data flow of mobile phone-centric sensing for user context recognition.

1.2. User Involvement in the sensing process

Based upon the level of the 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 does not require participation. While the identified system components introduced in the previous section essentially apply for both of these categories, the techniques for realising 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 compared to an opportunistic approach, significantly reduces the sensing, calibration and classification challenges. 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 (e.g. time, duration, location, space, contents, etc.) are dependent upon a participant's enthusiasm and willingness to collect data during their daily life. Moreover, collected data are affected by a bias of the user's knowledge/opinion during the data collection. The problematic effects of this fact are well known and carefully considered in data collection methods for human subject studies [Mcniell 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 main challenges in opportunistic sensing systems is determining how to transfer the required sensing functionality and intelligence to mobile phones without jeopardizing the phone experience caused by the additional processing overhead. For example, the position of the phone relative to 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. 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 realise [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 .2 in appendix .1 shows how they have been utilised in different mobile phone centric 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 accelerometers 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 been long 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. [Judd and Levi, 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, activity level(e.g. [Redmond and Hegge, 1985], [Bouten et al., 1997], [Wu et al., 2007], [Choudhury and Consolvo, 2008] and [Lester et al., 2006]) and ambulatory monitoring (for extensive discussion in this field refers to [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] and [Choudhury and Consolvo, 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 smart phones. Their primary purpose is to detect the changes in the orientation of the mobile phone so as to rotate the screen's display in accordance with the phone's orientation. Recent studies have utilized 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. A comparison between the required capabilities for activity recognition and the characteristics of the embedded accelerometers in current off-the-shelf mobile phones will clarify this issue in the following.

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 the vertical acceleration being the most dominant component, it is not advisable to neglect the horizontal acceleration [Lafortune, 1991]. In [Mathie et al., 2004] a comprehensive analysis of acceleration measurements with respect to different daily activities is provided. 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 acceleration magnitudes. For many researchers, the detection of walking activity as the most frequent daily activity [Kunze et al., 2005] is of great importance.

Studies about the range of magnitude and the frequency range of acceleration generated from body during daily life activities (e.g. [Antonsson and Mann, 1985] , [Aminian et al., 1995] and [Bouten et al., 1997]) confirm that the capabilities provided by current mobile phone embedded accelerometers are sufficient for detection of almost the same range of activities as with the current wearable approaches. For example, Cappozzo [Cappozzo, 1989], has stated that during walking, upper body accelerations in the vertical direction have been found to vary from −0.3 to 0.8g or Hilla and Sun [Hilla and Sun, 1993] have found that the major energy production 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) that are embedded in Nokia and Apple mart-phones ([Yang, 2009], [Hailes et al., 2009]) provide a 3-axial measurement with a configurable range of ± 2 g or ± 8 g and output rate of 100 Hz or 400 Hz [Anon., 2008].

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 are practically often not achievable on mobile phone due to implementation constraints and access restriction to the full resolution and sampling frequencies by the corresponding APIs [Brezmes et al., 2009]. For example**,** experiments of the authors 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 frequency of 36Hz on a Nokia N95 device. The power consumption of accelerometers is very small compared to other sensing modalities. For instance, the LIS302DL energy consumption is less than 1 mW .

2.1.2. Gyroscopes.

Recently Apple (with the iPhone 4G), HTC and ST-Ericsson, have announced that their next mobile will be equipped with MEMS Gyro sensors[2]. 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 for digital camera stabilizing techniques (e.g. [Yong-xiang et al., 2009]). With a 100% market penetration in any camera that has more than 5 megapixels, camera stabilization is predicted to be the primary task of gyros in new mobile phones[3]. MEMS-based gyroscopes are believed 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, Wi-Fi 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 the user soundings. 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.

[2] http://www.eetimes.com/news/latest/showArticle.jhtml?articleID=224701537
[3] http://www.eetimes.com/news/latest/showArticle.jhtml?articleID=224701537

Bluetooth is a universal, low-cost interface for ad-hoc wireless connectivity initially developed by Ericsson in 1994, released in 1998 to operate in 2.4-2.48 GHz, and ratified as the IEEE standard 802.15.1. Bluetooth is designed for short range communication. For example the power class 2 of Bluetooth devices which is typically implemented in handsets is limited to approximately 10 meter with a transmission power around 2.45mW(4dbm) The version 2.0 (2004) of Bluetooth communication is capable of transmitting up to 3Mbit/s [Schiller, 2003]. The main application of Bluetooth for sensing purposes has been in logging local devices and communicating with external sensors or services. Every Bluetooth device is capable of performing device discovery so as to obtain information about other devices in their vicinity. This information includes the Bluetooth MAC address, which is also referred to as a Bluetooth identifier (BTID), device name and device type. The BTID is a 48 bit number which is unique for a particular 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 and 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 proximate 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 analyse 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 is suggested in [Eagle et al., 2009], the only option up to 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 belong to the same region but from different network providers. Sometimes this redundancy in data is filtered by locking the logging software by the Subscriber Identity Module (SIM) card provider (e.g. [Sohn et al., 2006]) or clustering towers based on LAC (e.g. [Anderson and Muller, 2006]). 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* and typically occurs when the received signals on a mobile phone drop below a pre-determined 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 (e.g. fluctuation of cell IDs in a metropolitan area may have different patterns as compared with 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 and velocity of the user nearly anywhere on earth. GPS is based on simultaneous propagation measurements 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 distance of the mobile phone and each of a number of satellites [Mishra, 2004] in two dimensions (latitude, longitudinal), when the receiver is able to see at least three satellites. Zhao outlines in his study [Zhao, 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 Time of Arrival (TOA) of these signals. The 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 later introduced) 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 [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 may emit during the trip. Many researchers have especially emphasized the unique opportunity that the use of mobile phone GPS

sensors can provide for studying the travelling behaviour 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 travelling behaviour [Ohmori et al., 2005]. Travelling information from mobile devices is used in a variety of applications such as traffic estimation [Herrera et al., 2010] or helping riders for navigation and driving tips [Barbeau et al., 2010].
Despite the high accuracy of GPS technology for outdoor localization, GPS is usually considered as the most power hungry localization technique for mobile computing [Gaonkar et al., 2008].

2.2.4. Wi-Fi.
IEEE 802.11 (Wi-Fi) 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 United States and Europe, while in Japan it ranges from 2.471 to 2.497 GHz. As compared to Bluetooth, the other available short rage wireless communication method, Wi-Fi provides communication ranges of up to 100 meters but with much higher power consumption (30-100mW). Wi-Fi connections can also provide higher rates (up to few hundred Mb/s for one-way data) and they have less limitations in terms of the maximum number of devices in a basic cell (unlimited in ad hoc mode and up to 2007 nodes in infrastructure mode). A comprehensive comparison between Bluetooth and Wi-Fi communication and protocols is provided in [Ferro and Potorti, 2005]. A Wi-Fi 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 Wi-Fi signal transmission range, the positioning accuracy is not sufficient and so supplementary information such as signal strength (e.g. [Krumm and Horvitz, 2004]) or signal triangulation and fingerprinting when multiple access points (e.g. [Kansal and Zhao, 2007]) or a combination of them (e.g. [Cheng et al., 2005]) have been utilized is required. A comparison between GPS, Wi-Fi, AGPS and GSM localization in [Gaonkar et al., 2008], has shown that after GPS, localization techniques based upon the detection of Wi-Fi access points is the most power demanding approach. As a result, Wi-Fi is typically used as a secondary and complementary instrument while in the company of Bluetooth [Miluzzo et al., 2008] or GSM (e.g. [Gaonkar et al., 2008]) signals 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 acts as their carrier. In this section we discuss sensors that can be used for sensing the surroundings of a user, such as a camera, magnetometer and microphone. Exploiting their capabilities for sensing environmental state, some researchers have utilised 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. Therefore, it is 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 (i.e. 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 [Takacs et al., 2008] in providing 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 back end server. Sometimes picture frames are used directly with no further processing (e.g. [Miluzzo et al., 2008] , [Larsen and Luniewski, 2009]) or utilise simple and computationally affordable techniques directly on the mobile phone [Wagner et al., 2010] (for a comprehensive discussion the reader is referred to [Gu et al., 2008]).
Opportunistic sensing with a camera is not as straight forward as it is with the aforementioned sensors. For instance, since the pictures are not taken intimately by a user, 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, adequate data management techniques are essential for resource constraint 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 utilise the Hall Effect[4]. The sensed magnetic field is a combination of earth magnetic field and surrounding objects. The effect of surrounding objects can be divided to deterministic interference including the effect of ferrous materials (soft iron) and magnetized materials (hard iron) and non deterministic interference. The effect of surrounding objects can distort or even superimpose the direction earth magnetic field in navigation proposes. In this case additional systems are required along magnetometers for compensation of surrounding interference. Currently, magnetometers that embedded in mobile phones are typically equipped with a Dynamic Offset Estimation (DOE) system for compensating 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 [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 gyro sensors are used for gait recognition. A similar approach has been implemented by [Lee and Mase, 2001] and [Lee and Mase, 2002]. Some personal navigation systems for mobile devices combine compass and GPS information. Assuming the user is aware of the mobile phone orientation, the system simply provides a comparison between the phone orientation (or the sensitive access of magnetometer) and the static directions (e.g. North, South...) of a map. Other examples of smart phone-based applications exploiting embedded magnetometers include a three degrees of freedom controller in 3D object rotation tasks based upon innovative techniques such as those proposed in [Katzakis and Hori, 2009]. Mobile phone embedded magnetometers 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, 460 µA of average current with measurements at 100 ms intervals [6 Axis Electronic Compass Chip - AK8976A - Asahi Kasei, 2006].

2.3.2. 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. Since 1876, when Emile Berliner invented the first microphone, it has been extensively used for converting analogue sound waves into electrical signals. As the initial idea for developing mobile phones was to ease the transmission and reception of voice, microphones have always been a part of mobile phones. 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 Choudhury and Consolvo have discussed in [Choudhury and Consolvo, 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 non-verbal features from the recorded sound before sharing any information. Non-verbal vocal cues such as the pattern of silent moments, pich 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 behaviour on his mobile phone. Another example of non-verbal 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. PRE-PROCESSING

In order to reduce data redundancy, noise and jitter in instantaneous sensor readings, measured values are usually passed to a pre-processing stage. Figure 2 provides a flow chart of a typical pre-processing stage of mobile phone-centric sensing systems. The pre-processing 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 immense, even when limited to the aforementioned sensors. There the discussion on 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*

[4] e.g. HTC Dream uses AK8976A Hall effect magnetometer from Asahi Kasei Micro devices [Katzakis and Hori, 2009]

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. It should be noted that despite the discussed problems, which affect all sensing systems that are developed on mobile phones, errors related to the specification of implemented mobile platforms such as added error 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.

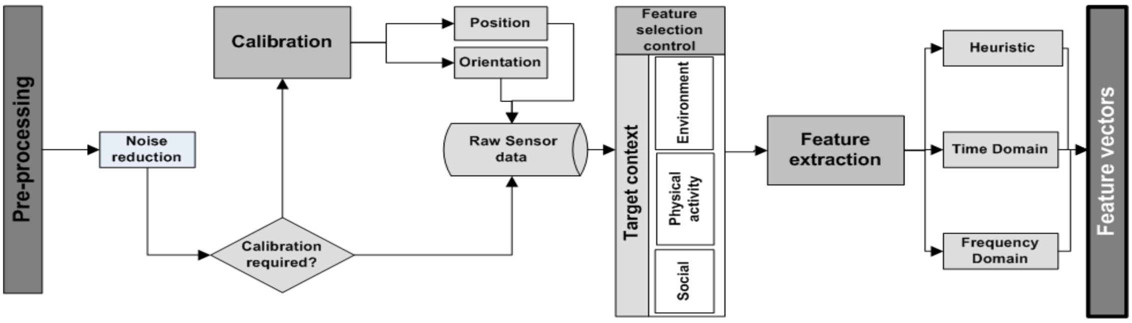


Figure 2. Overview of different procedures during the pre-processing stage

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, the mobile phones are carried and used in ways that are difficult to anticipate in advance for a particular user. Moreover, the output 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 can also help to provide more informative features[5]. Therefore, providing a pervasive sensing system on a mobile phone requires a calibration process to transfer the measured data into a 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 to identify adequate positions and confining the sampling to the moments that the mobile is in those. For instance, the data collection technique in [Miluzzo et al., 2008] takes photos when the user touches a key on the phone or in [Azizyan et al., 2009] a photo is taken when 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 caused by 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] and [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 device in a particular position. This makes the resolution of disorientation errors easier to accomplish and requires a minimum amount of 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 part of the section.

[5] For example when colours in an arbiter picture from the environment are used for user localization [Ofstad et al., 2008], information about the orientation of the phone can determine whether the colours belong to the ceiling or floor

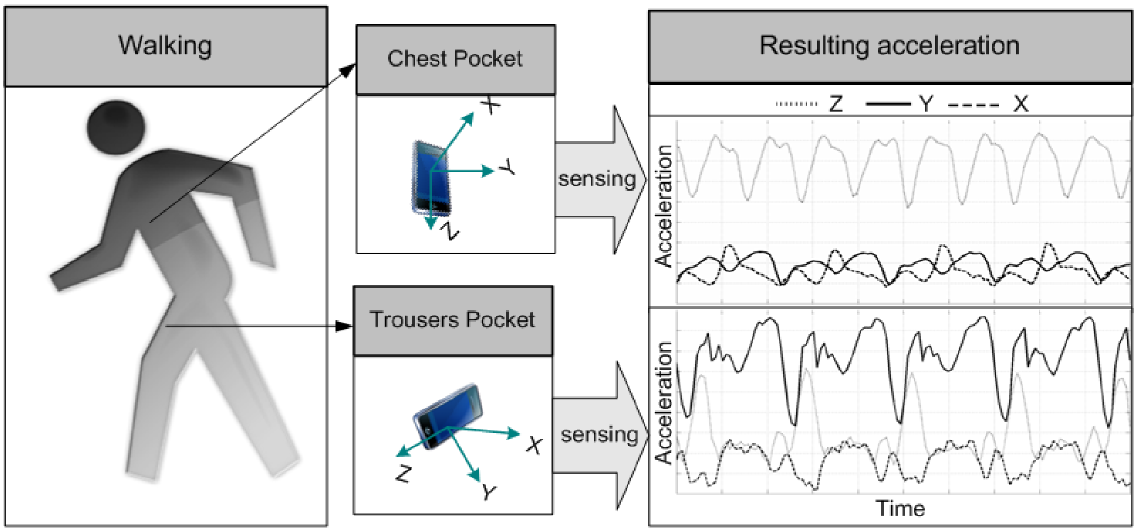


Figure 3. Variation in accelerometer orientation and position affects the measured acceleration pattern. Magnitude, frequency of the components and the axis of major components differs based upon the sensors' relative position and orientation relative to the user.

3.1.1. The effects of device position

The dependency between magnitude and the frequency of measured acceleration to the position of accelerometer on user's body was already 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 about 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, which subsequently try to directly detect the context regardless of a mobile's position. These methods usually require large databases and are less accurate as compared to other models. However classification is achieved more quickly. Calibration is not required with this method because all of the possible positions are predefined. In the 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 generalised data from different locations has shown that a reasonable accuracy can be achieved regardless of phone's position. However, the accuracy of the model increases significantly with the number of individual training data sets. The authors 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 the previous 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 during an activity changes, 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], the accelerometer signals during walking are used 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 it is a very frequent activity in everyday life. Examining several positions on the body, such as wrist, head, trousers' pocket and 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 [Kunze and Lukowicz, 2007], have then 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 of peaks along the three axis). Kawahara et al. [Kawahara et al., 2007] have exploited the unique behaviour of accelerometer signals in multiple situations so as 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 the chest pocket, accelerometer patterns caused by a person when stooping forward in the chair were used. For the trouser pocket position, fluctuations of the tilt angle during walking when were utilised. Furthermore the variance of the signals provided sufficient clues to determine when the phone was not with the user. Finally, Miuzzo has exploited vocal signals from mobile phone's microphone to extract a set of required features to estimate the device position. Here a sophisticated classification algorithm (Gaussian Mixture Model with 20 components) is adopted to classify the position of the device as in the pocket 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 the 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 positions near the Centre of Gravity (COG) of the subject (see, for example, [Mayagoitia et al., 2002], [Sekine et al., 2002], [Evans et al., 1991]) as suitable positions. A study in [Murray, 1967] shows that the applied force near the COG of the human body while walking is almost deterministic and undisturbed by individual characteristics[6]. The human centre of gravity , also referred to as the body's centre of mass, is located within the pelvic region in a standing position [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. In [Ichikawa et al., 2005], the researchers report 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] and [Ofstad et al., 2008]). For example, Bao and Intille [Bao and Intille, 2004] have investigated the effect of sensor position on mobile-centric activity recognition and suggested that positions near the hips are ideal positions. Inspired by Bao's findings, Miluzzo et al.'s study about different aspects of a mobile user's behaviour [Miluzzo et al., 2008] has encouraged the participants to place their mobile phones in their front or back trouser pocket.

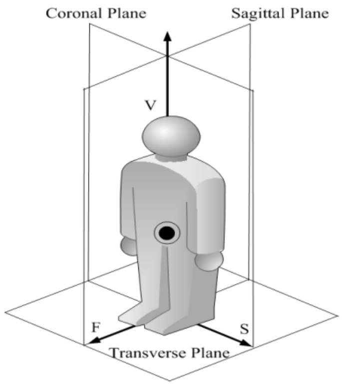


Figure 4. Demonstration of a body's coordination system and rotation planes. The intersection of the planes shows the position of CoG.

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 (i.e. omit the directional data) (e.g. [Gyorbiro et al., 2009], [Yang, 2009], [Santos et al., 2010], [Brezmes et al., 2009], [Kwapisz et al., 2010] and [Fleury et al., 2009]). However, such techniques discard the valuable information that sensing in multi-dimensions could provide. Rather, some studies have developed some calibration techniques to have higher dimensional 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

[6] . This fact has been also utilized 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])

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 the gravitational accelerations 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. In [Luinge et al., 1999] and [Kourogi and Kuratta, 2003], the gyroscope measurements have been used for determining the device's orientation. Its orientation is calculated by passing accelerometer and gyroscopes measurement values through a Kalman Filter. According to [Zhang et al., 2008], the processing of 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, the authors in [Zhang et al., 2008] 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 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 a 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 in [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 in this study has been successfully implemented in a mobile device (in particular a Nokia 810). A novel semi-analytical approach has been recently presented in [Hoseinitabatabaei et al., 2011] where the dynamic model of the movement of the body segments corresponding to the position of the device is used for recognizing the coordinate of the user. The model has shown to outperform the conventional PCA and GPS based approaches. Combining these techniques with the vertical direction identification provided 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 .1 summarizes the mentioned techniques along appropriate examples of available approaches. Having all the sensor data from the predefined references after the calibration process, the next step in pre-processing is to extract features from the calibrated data.

Table.1. 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 a 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, which affects both computation and 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. Firstly, 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. It is worth reminding the reader that 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 characterise 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 pre-processing 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 the 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 in back-end servers with no limitation in processing time, memory and energy consumption is available. 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 reduce power consumption, learning and classification problems [Choudhury and Consolvo, 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 in [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 (e.g. walking, running, etc.) (ca. 1 hour). In an effort to observe at least a subset of these fragments, many studies have exploited the 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 between a variety of daily physical activities, the position-based method can distinguish between different modes of movement. Accelerometers are especially considered to provide the most discriminative information for activity recognition [Choudhury and Consolvo, 2008], [Lester et al., 2006]. Respectively, 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]) sometimes additionally climbing (e.g. [Kwapisz et al., 2010]), cycling or driving [Bieber et al., 2009], [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 and Consolvo, 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] and [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 is used in distinguishing standing from sitting [Yang, 2009] [Miluzzo et al., 2008]. The signal variance is also utilised as a natural choice for estimating the intensity of activity. For example, [Ermes et al., 2008] has 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 [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 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 the above accelerometer-based features, logging the pattern of user locations over time is often sufficient to detect the user's activity motion. Consequently, all the sensing systems that are introduced for localization techniques, in principle, are able to provide such information about the user. However, the recognition level 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 has been a conventional source for inferring different states of user motion (e.g. [Sohn et al., 2006] and [Anderson and Muller, 2006]). By means of 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, 2006] and [Anderson et al., 2007] the change in *the number of unique LACs* along with the fluctuation of signal strength and the rate of change of cells has been used for identifying different modes of movement. GPS is also widely used for detecting movement. In [Miluzzo et al., 2008], the *GPS positioning information* over time is used for inferring the user's mode of movement such as being in a vehicle, running or stationary by estimating their speed. It is worth noting that since the activity recognition with localization techniques requires a comparison of the several subsequent locations of user, typically these techniques require a greater amount of time to determine the state of the user than systems that take advantage of inertial sensors.

Frequency domain features. Because of 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 cosine of 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 this type of 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 a gyroscope (measurements (e.g. [Lee and Mase, 2002]).

3.2.2. Features used for detecting social interactions.

Perceiving social signals by mobile phones to attain insight into one's daily social interactions has gained the attention of various researchers. Social signals refer to the non-verbal behaviours 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 main clue for having a social interaction.

Time domain features. In order to determine the presence of a social interaction as the first and foremost step for understanding social interactions, a number of techniques have been proposed. Lu et al. in the SoundSense project [Lu et al., 2009] have 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 [Lockheed and Nashua, 1996]. Calculating the low energy frame rate is also relevant since human conversations have more moments of silence than music and ambient noise [Lockheed and Nashua, 1996].

The physical and non-verbal behaviour of individuals conveys a significant amount of information about their behaviour in social interactions. A study by Viniciarelli [Viniciarelli et al., 2009], has identified the most important features of vocal and non-verbal behaviour as voice *quality*, *turn talking* and *silence/pauses* during speaking. These features can be extracted with a simple microphone, without directly analysing the user's speech. Such information's 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 direct feedback to the mobile phone user. The samples obtained from the accelerometers are also used to understand user social interactions. In [Kim et al., 2008], *average of body movements* within a fixed unit of time during a conversation is proposed to help the analysis of behaviour (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. in [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 differentiate the conversation moments from ambient noise. Researchers in [Lu et al., 2009] have introduced and implemented a number of frequency domain features for differentiating the human voice from music as well as 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. Another feature is *Spectral Roll-off Frequency (SRF)*, which is calculated as the 95th percentile of power distribution [Scheirer and Slaney, 1997]. A larger number of high frequency components in music compared to human voice lead 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. A further feature, namely the normalized weighted phase deviation as introduced in [Dixon, 2006], is determined by a weighting of the phase deviation of frequency bins in the spectrum by their magnitude. Ambient sound and music have less phase deviation than the human voice. Finally, *Relative Spectral Entropy (RSE)*, which 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 differentiate the human speech form other sounds.

Heuristic features. Bluetooth scanning is the most popular technique for detecting social interactions. Bluetooth scanning makes use of periodic invocations of the Bluetooth device discovery function in order to determine the devices (and other users) in proximity of the user. The presence of another user in proximity is considered as a potential social interaction. The technique exploits the uniqueness of the BTID which is transmitted by mobile phones together with Bluetooth-personal area network capabilities when queried. The logged BTID in [Miluzzo et al., 2008] is compared with a database of MAC addresses in order to infer whether a user is proximate to his/her friends. The proximity information is further used for identifying persons in ones vicinity, in order to subsequently establish correlations between people with the same application and calculating social status metrics (e.g. being popular). A case study by Eagle and Pentland [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 applications 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 are reported to a back to an end server. The collected data is then analysed to extract patterns of social relations thereby demonstrating 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 but 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] and [Bar-Noy and Kessler, 1993]) to infer the user's location and overlay it onto a map using a Geographic Information System (GIS). The inertial sensors such as accelerometers and gyroscopes are also utilised to detect the user's movement pattern in a known topology (i.e. dead reckoning) (e.g. [Blanke and Schiele, 2008][7], [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 explained by the reliance on absolute positioning systems, which usually do not require an analysis over time (or frequency).

---

[7]Relative positioning is used in contrast to absolute positioning systems such as GPS and UWB

Time domain features. Probability Density Functions (PDF) of the locations of cell towers 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, in [Ofstad et al., 2008] the *percentage of time that a user is in a standing state* is used for localization (e.g. being in a coffee shop or shopping centre), 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 conventional definition can be regarded as a measure of the flatness of a FFT spectrum. While ambient noise has a limited spectrum, music is typically spread across a wider range of frequencies. Other features worth mentioning are *Mel Frequency 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 features extraction is a computationally demanding process, they have been effectively used 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 recognising a user's environment make use of features from camera pictures, such as *illumination* (e.g. [Azizyan et al., 2009]) or the *colours* (e.g. [Ofstad et al., 2008] and [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 is 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. [Azizyan 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 in [Santos et al., 2010]. The noise level indicates in another study whether the user is attending a party [Miluzzo et al., 2008]. Here 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, 2006], [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 labelled traces of GPS. Another approach uses a static Bluetooth beacon [Eagle et al., 2009] or Wi-Fi ([Miluzzo et al., 2008] transmitters, to detect the presence of a user in a pre-determined location. Here the reception of signals from several transmitters each with a particular MAC addresses 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] and [Laasonen et al., 2004]). For instance, in [Eagle et al., 2009] the GSM signal strength has been logged on mobile devices in order to determine the cell towers in the vicinity and consequently the location of device.

3.2.4. Summary.
In this section, we have introduced the recent 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 in [Iso and Ymazaki, 2006], frequency and time domain-based features from sensor data have less 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 context recognition 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 from stairs climbing for healthcare approaches [Sekine et al., 2000]. However, in case of mobile phones where in contrast to wearable sensors the computational resources are utilised concurrently for a variety of different tasks, the realization of these computationally demanding features have been primarily limited to offline modes in the past (e.g. [Iso and Ymazaki, 2006]). The recently emerging powerful microprocessors for mobile phones make however the utilisation of time-frequency based features on mobile phones now 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 [Wittke et al., 2009] has used the Harr-like features that are computed similar to coefficients in Harr wavelet transform 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.

Different 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 sub actions, body posture or relative position of user and objects. Being driven from body model these features are more robust and less variant to variability of performing different activities [Zinnen et al., 2009]. For example, Zinnen in [Zinnen et al., 2009] have utilized various motion primitives such as moving the hand up or down or turning the hands along 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. In addition, their heavy computational burden has still remains an issue for implementation on mobile devices. Table.2 in appendix 1 comprises the type of features that have been utilized 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 Markove Models (HMM). However, the classification methods that are implemented on handheld devices must be adapted to the limitation in computational capabilities 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 context inference process is a further distinctive parameter.

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 as solutions are not given [Pietquin, n.d.]. Selecting each type of learning model affects the design of the labelling 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:

$$MSE = Noise^2 + Bias(f(x))^2 + Var(f(x)). \quad (1)$$

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 (1) *Noise*, represent 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 which unfortunately is not possible due to the natural bias-variance trade off. For example, while a learning model may suffer from under-fitting problems (high bias error) due to training on very large data sets, it is also susceptible to over-fitting (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 data set 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 for 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 as compared to similar studies with a small set of users such as reported in [Yang, 2009]. One of the main drawbacks of these approaches is the estimation of the number of required participants to have adequate varieties in the database for training. 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 an online learning (e.g. [Könönena et al., 2010] or [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 improve using crowd-sourced labelled data for training, while the probable mislabelling errors (e.g. humanistic errors) are addressed by utilizing the data similarity. A combination of community guided learning and active learning is proposed by Berchtold [Berchtold et al., 2010] where a service-based recognition architecture is utilized for recognizing user physical activities. Here on a backend server a Global Trainer Service 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 classifiers modules on the device. A different approach is to use features that do not change significantly among different users during learning process [Kawahara et al., 2007] .
The learning techniques are determined according to the classification technique of the choice. In the next section, different classification techniques used in mobile centric applications are introduced.

4.2. Classification techniques

As discussed by Ye in [Ye, 2004], almost all the classification algorithms are solving 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. It 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), *Maximum a Posteriori* (MAP) or *mean posteriori* techniques usually 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 back end servers, our emphasis for a mobile-centric sensing system is on classification techniques that have been implemented on mobile devices. Figure 5 shows taxonomy of the algorithms that will now be presented, analysing the recent approaches in developing classification algorithms on mobile phones.

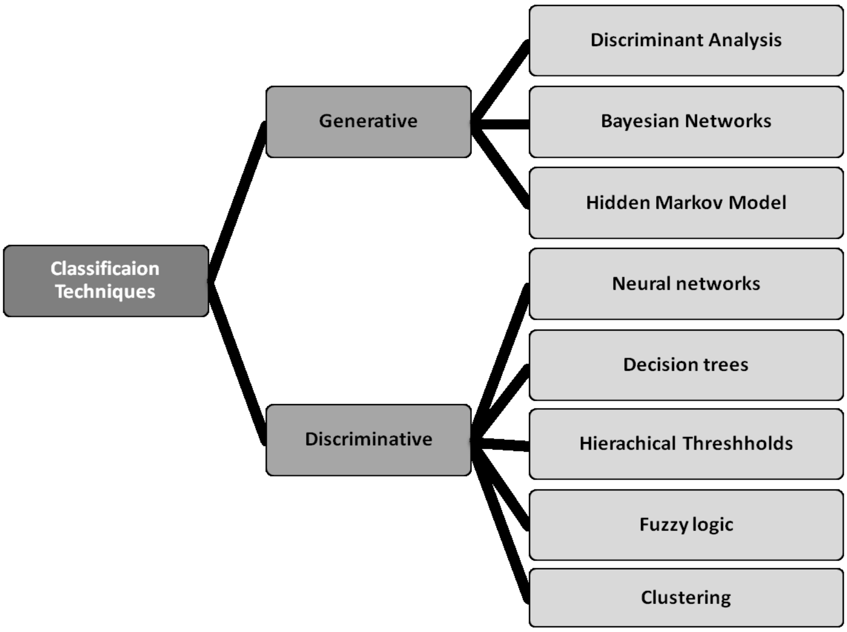


Figure 5. A taxonomy of the classification techniques that have been successfully implemented for context recognition

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 over-fitting (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 leafs 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 J48 decision tree algorithm, used in user social context recognition and 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 reported in [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 has 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 and 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 over-fitting problems([Blum et al., 2006], [Santos et al., 2009]) and cannot be used for generic applications unless large data sets are available for training.

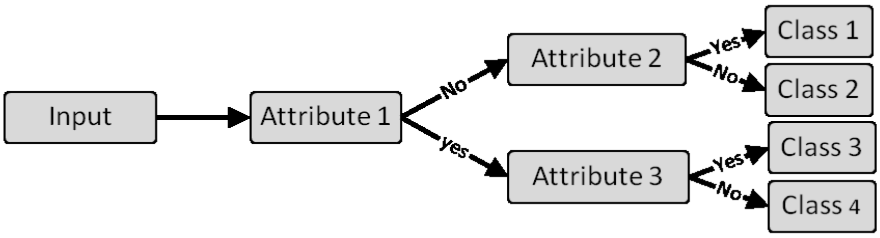


Figure 6. Decision Tree: structure of a decision tree with 3 attributes which classifies the input into 4 different classes

4.2.1.2. Neural Networks. The work undertaken on artificial neural networks is motivated from complex, non-linear 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 number of hidden neurons, is found by comparing the performance of classifiers with different number of hidden neurons. A feed forward network with 3 hidden states is depicted in Figure 7.

Bruns [Bruns et al., 2007] has 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 in [Anderson et al., 2007] has implemented a neural network with 8 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, the 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 recognised activity. Similar to the decision trees, training neural network is usually considered computationally expensive and consequently performed offline (e.g. [Gyorbiro et al., 2009] [Anderson and Muller, 2006]). Therefore the neural networks are not the suggested when the system is subject to frequent retraining.

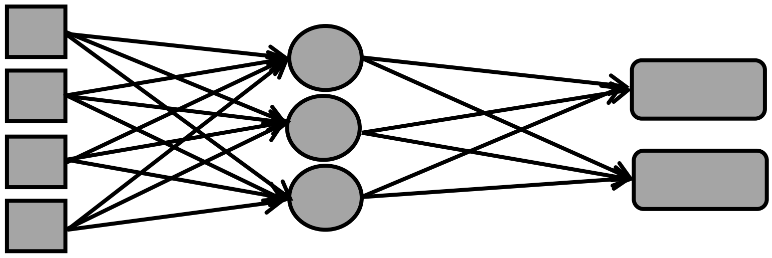


Figure 7. Structure of a feed-forward neural network with four inputs, three hidden states and two classes of outputs

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 over-fitting susceptibility. Additionally, its 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 the 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 instance, in [Haykin, 2009], fuzzy logic is used 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 in [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[8] which require labelled data during training, some studies have used clustering as unsupervised leaning algorithms for both classification (e.g. [Brezmes et al., 2009]) and calibration (e.g. [Anderson and Muller, 2006]). The clustering is described as an unsupervised classification of patterns (observation, data items, or feature vectors) into groups of clusters. [Jain and Murty, 1999] . For an extensive discussion about different clustering techniques, refer to [Jain and Murty, 1999].

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

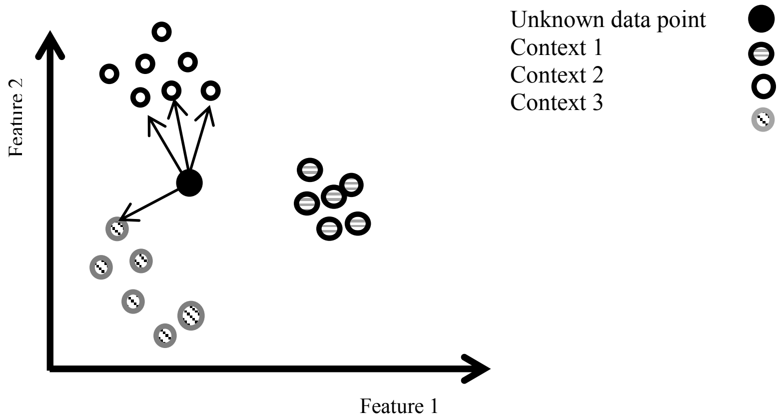


Figure 8. KNN clustering: The input is the unknown data and its four nearest neighbours. The clustering is performed in a two-dimensional feature space.

The activity recognition technique in [Brezmes et al., 2009], has used the K nearest approach which is trained based upon user preferred 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. In [Mirkin, 2005] Mirkin has described the *K-means* algorithm as a major clustering technique which is fast and straightforward. Based upon this technique a multi-dimensional space of features is divided into *K* clusters through a recursive algorithm of finding 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 faced by memory constraints. In addition, the K-means algorithm is susceptible to local minima's 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 magnitude of the mean and standard deviation of accelerometer signals are divided into six clusters. The clustered data are then labelled based upon the distance between their corresponding centroids to different classes of decision trees. It is however unclear as to whether the algorithm has been actually implemented on a mobile phone.

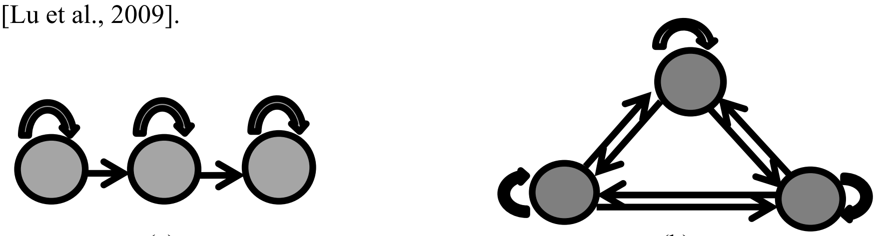
4.2.2. Generative models.

Generative analysis such as a *Hidden Markov Model* (HMM) or its hierarchical extensions demonstrate significant potential for the classification of everyday activities. However there is a significant challenge for

[8] The learning process for artificial neural networks can be both supervised and unsupervised.

porting resource-intensive HMMs to a mobile device. As a generative model which does not involve many mathematical calculations, discrete HMM has widely been used for smoothing the classification results by finding the most probable output, considering one or number of previous states [Wu et al., 2007], [He et al., 2007]. For a detailed discussion of related issues the reader is referred to [Attalah 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 as 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. Cappe et al. [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 9a shows a left-to-right HMM, where the Markov chain starts in a particular state and after number of transitions 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 9b, 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 discrete HMM.
In their work [Anderson and Muller, 2006] [Anderson et al., 2007] [Anderson and Muller, 2006], Anderson et al. have implemented a discrete HMM model on mobile phones for recognizing user activity from GSM signals. Here, the observation data is based on signal strength fluctuation and cell fluctuations, which is mapped onto a set of fifteen 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 have been also 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].



(a) (b)

Figure 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.

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

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

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 be also 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-dimension as:

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

Where, the subscript k indicates the class and μ is the mean vector $\mu \in R^n$and Σ is covariance matrix $\Sigma \in R^n * R^n$.
Blume et.al in [Blum 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. The model is claimed to be faster than HMM while providing comparable results and is also immune to over-fitting problems in contrast to 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 data sets 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 in [Ofstad et al., 2008], has used KDE during implementing 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 [Cho et al., 2007] have exploited modular Bayesian networks to recognize relevant or novel landmarks during movement in daily life and 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, 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 the user contexts. Here, the performance of a classifier in recognizing the demanded context determines the classifier of the 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 about 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 et.al [Caruana and 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 evaluating data set 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 applications 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). Meaning the priority is with the classifications that minimize the TN. These situations the cost of making different error should be taken in to account. Cost sensitive classification or cost-sensitive learning techniques are different techniques for incorporating the cost into classifiers. Through the former technique the costs are ignored during learning and are then applied to

predictions as a cost matrix and the later one considers the costs during the learning procedure. The required cost matrix in the first technique and the weights (costs) in the second one 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 penalizing the model complexity and minimizing the following expression.

$$-2\log L + P \tag{4}$$

Where L is the maximum likelihood and P is the penalty for complexity. 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 divided into voting, stacking and cascading. Ravi et al. claim in [Ravi et al., 2005] that the performance of base-level classifiers for activity recognition can be improved by using Plurality voting technique. However, the real-time implementation of these techniques remains an open research issue.

4.3. Labelling

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 labelled before training. There are various ways to perform labelling 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 labelling 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 thereby degrade the classification technique in reality [Azizyan et al., 2009]. Due to the dependency on hand-labelled data for training classifiers, applications that exploit these techniques are typically constrained to a small set of pre-defined aspects of user context (classes). Accordingly, a more challenging approach is to perform online learning and labelling in order to scale the available classes to a larger number of distinctive classes as required. Current efforts for labelling the new events that have not been covered during initial training utilize the user's intelligence when an unknown context is recognized. For example, in [Lu et al., 2009] a hybrid approach for supervised and unsupervised learning is proposed, 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 authorised to add (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 considerations in selecting the functionalities and algorithms can both fulfil the application requirement 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 process 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, mean, intensity and number of peaks in a window of samples seem to be essential inputs to infer user physical activity. The most distinctive and informative features available for determining user social interactions are the user's proximity and vocal behaviour. Finally, user environment can be characterized by combining absolute positioning data with heuristic features such as colour or typical user behaviour such as a location fingerprint.

Selecting a proper context recognition technique is one of the challenges that still require further addressing. Before selecting a classification technique, an appropriate strategy for training and labelling 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 computational efficient but requires a careful consideration about the generality of the training data set in order to avoid over-fitting 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. The 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 labelling 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 case of the classification techniques, an initial review about the introduced classification methods demonstrates that the decision trees and the neural networks provide satisfactory results for most of the applications. In addition, in small network (or tree) sizes they can be easily trained and implemented on mobile devices. However, they are prone to over-fitting problems. Developing hierarchical thresholds for hierarchical approaches is time consuming. However similar to the decision trees; they 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 over-fitting 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 fulfil the requirement of most of applications. Inferring the context on the phone has been emphasized to provide a number of advantages [Miluzzo et al., 2008]. It present resilience to cellular or Wi-Fi 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 the monetary cost by merging consecutive phone uploads and also protects user privacy and the data integrity by keeping the raw data on the phone. Finally, it provides an opportunity for creating user labelled contexts.

When a two-stage model is used, the inferred context or the learned parameters from user behaviour can be provided to the backend servers for further processing. Especially in case of real-time sensing applications, uploading the data to a backend server may help to 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 analyzing 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, as described in [Gaonkar et al., 2008], energy consumption increases form GSM to Wi-Fi-based localization and GPS schemes while the accuracy decreases from GPS to Wi-Fi and GSM methodologies. As another example, updating data on 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 back-end server of a typical mobile phone device (e.g. Bluetooth, HTTP+3G, HTTP+ Wi-Fi); the battery level of the device, the energy cost of the connections along with the available data rate and connection coverage are the parameters needed to determine the connection of the 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 offload part of the data processing from the phone onto a backend server [Kanjo et al., 2009] to help to reduce the power consumption. The proper application of such methods leads to developing a power aware duty cycle for both sensing and uploading while the application responsiveness is not affected. The table.2 in Appendix.1 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, realising 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 the 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 SDKa (Software Development Kit) and APIs (Application Programming Interface). For example, it is currently difficult to establish a consistent sensing frequency that does not change with CPU load. Effective programming for managing the sensing process can, to some extent, mitigate the problem. In 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]. Finally, inspired from 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 [Mobile sensing group, n.d.]. Such methods would help to access to sensing data from other devices when the sensors are not available or 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 feature generation 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 because of 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, Kononena et.al [Könönena 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 compromise 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 ignoring the repetition in common processes, evaluating the required overall computational and storage costs for feature generation would be different in comparison with a linear addition of separate processes. 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. Therefore, the feature selection system must 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 (SFB/S), Sequential Floating Forward Selection (SFFS) in [Könönena et al., 2010] or the boosting-based technique in [Choudhury and Consolvo, 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, while applicable with mobile phones computational constrains, has remained a major challenge in this area.

6.3. Labelling

Distinguishing and labelling different contexts forms another major challenge. In the real world, drawing boundaries between different aspects of user behaviour is difficult. It is likely that people at home sometimes exhibit the same behaviour as they do in their office or even perform different activities at the same time. The complex social behaviour 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 behaviour appears to be essential for such systems in order to be useful in real-world situations. Another important shortcoming in current labelling 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 labelling from available clues in user context such as *common sense* reasoning [Havasi et al., 2009] seems 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 times series of activities and labelling 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 to determine 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. in [Kapadia et al., 2009] have envisioned some of the related security challenges in opportunistic sensing. The authors argue that the new characteristics of sensing architectures, including high mobility, opportunistic networking, strong but discontinues connectivity and relatively plentiful power in one hand and dealing with very personal information in other hand has posed new challenges for information security. These challenges cannot be addressed with previous security solutions such as cryptography and privacy-pre-serving 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 [Hakkila 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 solving the privacy issue appears to 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 user or/and 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 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 physical state and orientation (e.g. [Ghiani and Paternò, 2010]). An important role will play 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 the pattern or profiling the users for targeted advertisement or optimization of different service delivery to a user's the handset. Recognizing the

patterns and profiling different aspects of user behaviour is currently a very fast growing research area and includes different perspective of user behaviour. One interesting existing example is the study of routines of physical activities of elderly people in [Huynh and Schiele, 2008] where *topic modelling* technique is used to detect the patterns form longitudinal data that is generated via an activity recognition application. Another recent study profiles user's according to the dynamics of their social networks in [Candia et al., 2008] It is likely that in the near future the similar researches extends the current binderies to other aspects of user context and even to the of study the interconnections between different contexts.
In large-scale applications, network providers can take advantage of user context data for modelling user behaviour in order to manage their resources and service allocations more effectively. Environmental monitoring applications are another type of the emerging applications on mobile phones. Here each mobile phone acts as a sensor for monitoring particular parameter of the user environment (e.g. noise level [Santini et al., 2009]and CO2 footprints [Mun et al., 2009], 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 is envisioned to be monitored with these platforms phones. Currently health care applications can be easily extended from personal monitoring to large-scale monitoring for epidemiological purposes. Particularly, 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 analysing a spectrum of social networks ranging from personal and small groups to large-scale communities. The pervasive data entailing user behaviour 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 more likely in the near future that the use of mobile phones with pervasive sensing and social signal processing capabilities, share the current multi-million pound market of social surveys. Examples range from smaller scale studies such as organizational behaviour to [Cross et al., 2002]. to large-scale ones such as International Social Survey Programme [GESIS – Leibniz Institute for the Social Sciences, 2009] and European Social Survey [European Social Survey , 2009].

7. CONCLUSION

The 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 smart phones have become increasingly multi-purpose platform, it is still a challenging task to add opportunistic sensing and context processing capabilities, without jeopardizing the user's mobile phone experience.
A large diversity of different opportunistic sensing techniques and applications utilizing those have been developed in the recent years. This paper analysis these approaches and represents a first attempt to classify the utilised techniques and methodologies as distinct components of a mobile sensing architecture. The resulted architecture comprises three major stages, namely sensing, pre-processing and context recognition involving a variety of techniques to fulfil one or more tasks inside each of the stages. For each of the stages a thorough analysis of advantages and shortcoming of the currently implemented techniques has been provided. Ensuring an adequate quality of context while considering device constraints for user context-recognition requires a deeper understanding of these and how they interplay with each other. The paper has contributed towards this understanding by deriving recommendation from the former analysis on how these techniques at different stages of the architecture should be combined to build more reliable mobile sensing platforms.
The paper concludes with 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, labelling of complex and unknown events and the preservation of user privacy of the user 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.

8. REFERENCES

ALLEN, F.R., AMBIKAIRAJAH, E., LOVELL, N.H., AND CELLER, B.G. 2006. An adapted Gaussian mixture model approach to accelerometery-based movement classification using Time Domain Features. In *Proceedings of 28th IEEE EMBS Annual International Conference*, New York, 2006. IEEE. pp.3600-03.

AMINIAN, K., ROBERT, P., JEQUIER, E., AND SCHUTZ, Y. 1995. Incline, speed, and distance assessment during unconstrained walking. *Medicine and Science in Sports and Exercise*, 27(2), pp.226-34.

ANDERSON, I. et al. 2007. Shakra: Tracking and sharing daily activity levels with unaugmented mobile phones. *Mobile Networks and Applications*, 12(2), pp.185-99.

ANDERSON, I., AND MULLER, H., 2006. *Practical Activity Recognition using GSM Data*. CSTR-06-016 Technical Report. Department of Computer Science, University of Bristol.

ANDERSON, I., AND MULLER, H. 2006. Practical Context Awareness for GSM Cell Phones. In *International Symposium on Wearable Computing (ISWC)*, Montreux, 2006. pp.127-28.

Anon., 2008. [Online] STMicroelectronics Available at: http://www.st.com/stonline/products/literature/ds/12726/lis302dl.htm [Accessed 23 June 2010].

European Social Survey. 2009. [Online] Available at: http://www.europeansocialsurvey.org/.

GESIS – Leibniz Institute for the Social Sciences. 2009. [Online] Available at: http://www.gesis.org/en/services/data/survey-data/issp/modules-study-overview/social-networks/2001/.

ANTONSSON, E.K., AND MANN, R.W. 1985. The frequency content of gait. *Journal of biomechanics*, 18, pp.39-47.

AOKI, P.M. et al. 2002. Sotto Voce: exploring the interplay of conversation and mobile audio spaces.In *Proceedings* of *ACM SIGCHI Conference on Human Factors in Computing Systems*, Minneapolis, 2002. pp.431-38.

ARIKAWA, M., KONOMI, S., AND OHNISHI, K. 2007. NAVITIME: supporting pedestrian navigation in real world. *Pervasive Computing*, 7, pp.1536-268.

ASAHI, K., *6 Axis Electronic Compass Chip - AK8976A - Asahi Kasei*. 2006. [Online] Available at: http://embedded-system.net/6-axis-electronic-compass-chip-ak8976a-asahi-kasei.html [Accessed 1 July 2007].

ATALLAH, L. et al. 2009. Real-time activity classification using ambient and wearable sensors. *IEEE Transaction on Information Technology in Biomedicine*, 13(6), pp.1031-39.

ATTALAH, L., AND YANG, G. 2009. The use of pervasive sensing for behaviour profiling-a suvey. *Pervasive and Mobile Computing*, 5(5), pp.447-64.

AZIZYAN, M., CONSTANDACHE, L., AND CHOUDHURY, R.R. 2009. SurroundSense: Mobile phone localization via ambience fingerprinting. In *The 15th Annual International Conference on Mobile Computing and Networking, MobiCom'09*, Beijing, 2009. ACM. pp.261-71.

BAHL, P., AND PADMANADHAN, V.N. 2000. RADAR: An in-building rf-based user location and tracking system. In *proceedings of the Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2000*, Tel Aviv , 2000. pp. 775 -84.

BAO, L., AND INTILLE, S.S. 2004. Activity recognition form user-annotated acceleration data. In *In 2nd International Conference, PERVASIVE '04*, Vienna, 2004. pp.21-23.

BARBEAU, S.J. et al. 2010. Travel assistance device: utilizing global positioning system-enabled mobile phones to aid transit riders with special needs. *IET Intelligent Transport Systems*, 4, pp.12-23.

BAR-NOY, A., AND KESSLER, I. 1993. Tracking mobile users in wireless communications networks. In *Proceedings* of *twelfth annual joint conference of the IEEE computer and communications societies. networking*, San Francisco, 1993. pp.1232 - 1239.

BASU, S. 2003. A linked-HMM model for robust voicing and speech detection. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, Hong Kong, 2003. pp.816-19.

BERCHTOLD, M. et al. 2010. ActiServ: Activity recognition service for mobile phones. In *International Symposium on Wearable Computers (ISWC)*, Seoul, 2010. pp.1-8.

BHATTACHARYA, A., AND DAS, S.k. 1999. LeZi-update: an information-theoretic approach to track mobile users in PCS networks. In *Proceedings* of *the International Conference on Mobile Computing and Networking*, Seatel, 1999. pp.1-12.

BHATTACHARYA, A., MCCUTCHEON, E.P., SHAVARTZ, E., AND GREENLEAF, J.E. 1980. Body Acceleration distribution and O2 uptake in humans during running and jumping. *Journal of Applied Physiology*, 49, pp.881-87.

BIEBER, G., VOSKAMP, J., AND URBAN, B. 2009. Activity recognition for everyday life on mobile phone. In *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*. C. Stephanidis, Ed. Springer Berlin / Heidelberg,

BLANKE, U., AND SCHIELE, B. 2008. Sensing Location in the Pocket. In *10th International Conference on Ubiquitous Computing (UbiComp)*, Seoul, 2008. pp.21-24.

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

BOUTEN, C.V. et al. 1997. A triaxial accelerometer and portable data processing unit for assessment of daily physical activity. *IEEE Transaction in Biomedical Engineering*, 44(3), pp.136-47.

BREZMES, T., GORRICHO, J.L., AND COTRINA, J. 2009. Activity recognition form 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*, 2009. Springer. pp.796-99.

BRUNS, E., BROMBACH, B., ZEIDLER, T., AND BIMBER, O. 2007. Enabling mobile phones to support large-scale museum guidance. *IEEE MultiMedia*, 14(2), pp.15-25.

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

CANDIA, J. et al. 2008. Uncovering individual and collective human dynamics. *Journal of Physics A: Mathematical and Theoretical*, 41(22), pp.224015-26.

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. *Journal of Biomechanics*, 15, pp.599-609.

CARUANA, R., AND MIZIL, A.N. 2006. An empirical comparison of supervised learning algorithms. In proceedings of *the 23th International Conference on Machine Learning*, Pittsburgh, 2006.

CHENG, Y.C., CHAWATHE, Y., LAMARCA, A., AND KRUMM, J. 2005. Accuracy characterization for metropolitan-scale Wi-Fi localization. In *Proceedings of the Mobysis 2005*, NewYork, 2005. ACM. p.233–245.

CHEN, D.M. et al. 2009. Streaming mobile augmented reality on mobile phones. In *Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality*, Orlando, FL, 2009. IEEE Computer Society. pp.181-82.

CHO, S.B., KIM, K.J., HWANG, K.S., AND SONG, I.J. 2007. AniDiary: daily cartoon-style diary exploits Bayesian networks. *Journal of Pervasive Computing*, 6(3), pp.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 *Ninth International Workshop on Frontiers in Handwriting Recognition, IWFHR-9 2004*, Tokyo, 2004. pp.106-11.

CHOUDHURY, T., AND CONSOLVO, S. 2008. An embedded activity recognition system. *Pervasive Computing*, pp.32-41.

CRK, I., ALBAINALI, F., GNAIDY, C., AND HARTMAN, J. 2009. Understanding Energy Consumption of sensor Enabled Applications on Mobile phones. In *31st Annual International Conference of the IEEE EMBS*, Minneapolis, 2009. pp.6885-88.

CROSS, R., PARKER, A., AND BORGATTI, S.P. 2002. Making invisible work visible : using social network analysis to support strategic collaboration. *California Management Review*, 44(2), pp.25-46.

DARTMOUTH COLLEGE, *Mobile sensing group*. n.d. [Online] Available at: http://sensorlab.cs.dartmouth.edu/ [Accessed 28 October 2010].

DAS, T. et al. 2010. Prism: platform for remote sensing using smartphones. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*, San Francisco, 2010. pp.63-76.

DAVIS, M. et al. 2006. Using context and similarity for face and location identification. In *IS & T / SPIE Electronic Imaging Conference*, San Jose, CA, 2006. pp.60610E.1-.9.

DELIGNE, S. et al. 2002. A robust high accuracy speech recognition system for mobile applications. *IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING*, 8(10), pp.551-62.

DESELAERS, T., HEIGOLD, G., AND NEY, H. 2008. SVMs, Gaussian mixtures, and their generative/discriminative fusion. In *19th International Conference on Pattern Recognition, 2008. ICPR 2008.* , Tampa, 2008. pp.1-4.

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

DUDA, R.O., HART, P.E., AND STORK, D.G. 2000. *Pattern Classification*. 2nd ed. Wiley,

EAGLE, N., AND PENTLAND, A. 2005. Social Serendipity: Mobilizing social software. *Pervasive Computing*, pp.28-34.

EAGLE, N., AND PENTLAND, A. 2006. Reality Mining: Sensing complex social systems. *Personal Ubiquitous Computing*, 10(4), pp.255-68.

EAGLE, N., QUINN, J.A., AND CLAUSE, A. 2009. Methodologies for continious cellular tower data analysis. In *Proceedings of the 7th International Conference on Pervasive Computing ,Pervasive '09*, Nara, 2009. pp.342-53.

ERMES, M., PARKAA, J., AND CLUITMANS, L. 2008. Sensors, Advancing from offline to online activity recognition with wearable. In *30th Annual International IEEE EMBS Conference*, Vancouver, 2008. IEEE. pp.4451-54.

EROL, B., ANTÚNEZ, E., AND HULL, J.J. 2008. HOTPAPER: Multimedia interaction with paper using mobile phones. In *ACM Multimedia Conference*, Vancouver, 2008. ACM. pp.399-408.

EVANS, A.L., DUNCAN, G., AND GLICHRIST, W. 1991. Recognizing acceleration in body movements. *Medical and Biological Engineering and Computing*, 29(1), pp.102-04.

FANG, L. et al. 2005. Design of a wireless assisted pedestrian dead reckoning system- the NaveMote experience. *IEEE Transactions on instrumentation and measurements*, 54(6), pp.2342-58.

FAWCETT, T., AND PROVOST, F. 1996. Combining data mining and machine learning for effective user profiling. In *Proceedings of Knowledge Discovery and Data Mining 1996, KDD96*, Protland, 1996. AAAI. pp.8-13.

FERRO, E., AND POTORTI, F. 2005. Bluetooth and Wi-Fi wireless protocols: a survey and a comparison. *Wireless Communications, IEEE*, 12(1), pp.1536-284.

FLEURY, A., NOURY, N., AND VACHER, M. 2009. A Wavelet-Based pattern recognition algorithm to classify postural transitions in humans. In *Proceedings the 17th European Signal Processing Conference*, Glasgow, 2009. EURASIP. pp.24-28.

FLEURY, A., VACHER, M., AND NOURY, N. 2010. SVM-based multi-modal classification of activities of daily living in health smart homes: sensors, algorithms and first experimental results. *IEEE TRANSACTION ON INFORMATION TECHNOLOGY IN BIOMEDICINE*, 14(2), pp.274-83.

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

FRIGO, M. 1999. A fast fourier transform compiler. In Proceedings *of the 1999 ACM SIGPLAN Conference on Programming Language Design and Implementation*, Gorgia, 1999. pp.1-12.

FROEHLICH, J. et al. 2007. MyExperience: 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*, New York, 2007. ACM. pp.50-57.

GAONKAR, S., LI, J., CHOUDHURY, R.R., AND COX, L. 2008. Micro-blog: Sharing and querying content through mobile phone and social participation. In *Proceeding of the 6th international conference on Mobile systems, applications, and services*, 2008. ACM. pp.174-86.

GHIANI, G., AND PATERNÒ, F. 2010. Supporting mobile users in selecting target devices. *Journal of Universal Computer Science*, 16(15), pp.2019-37.

GOLDING, A., AND LESH, N. 1999. Indoor navigation using a diverse set of cheap, wearable sensors. In *Proccedings of International Symposium on Wearable Computers (ISWC99)*, Washington, DC, 1999. pp.26-36.

GONZALEZ, M.C., HIDALGO, C.A., AND BARABASI, A.L. 2008. Understanding individual human mobility patterns. *Nature*, 7196, pp.779-82.

GRISWORLD, B.G. et al. 2002. Using mobile technology to create opportunistic interaction on university campus. In Proceedings *of Ubicomp 2002 workshop on supporting Spontaneous Interaction in Ubiquitous Computing settings*, Munich, 2002.

GU, J., MUKUNDAN, R., AND BILLINGHURST, M. 2008. Developing mobile phone ar applications using J2ME. In *Image and Vision Computing New Zealand*, Christchurch, 2008. pp.1-6.

GYORBIRO, N., FABIAN, A., AND HOMANYI, G. 2009. An activity recognition system for mobile phones. *Mobile Networks and Applications*, 14, pp.82-91.

HAILES, S., SICARI, S. AND ROUSSOS, G., Eds. 2009. *Sensor Systems and softwears*. Springer,

HAKKILA, J., AND CHATFIELD, C. 2005. Its like if you opened someone else Letter- user percived privacy and social practices with SMS communication. In Proceedings *of the 7th international conference on Human computer interaction with mobile devices & services*, Salzburg, 2005. ACM. pp.219 - 222.

HARO, A., MORI, K., SETLUR, V., AND CAPIN, T. 2005. Mobile camera based adaptive viewing. In *ACM International Conference on Mobile Uibiquitous Multimedia*, Christchurch, 2005. pp.78-83.

HAVASI, C., PUSTEJOVSKY, J., SPEER, R., AND LIEBERMAN, H. 2009. Digital intuition: applying common sense using dimensionality reduction. *IEEE Intelligent Systems,* 24(4), pp.24-35.

HAYKIN, S.S. 2009. *Neural networks and learning machines*. Prentice Hall,

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, 2007. EMBS 2007.*, Lyon, 2007. IEEE. pp.3192-95.

HERRERA, J.C. et al. 2010. Evaluation of Traffic data obtained via GPS-enabeled mobile: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies*, 18(4), pp.568-83.

HIGHTOWER, J., AND BORRIELLO, G. 2001. Location systems for ubiquitous computing. *IEEE Computer Magazine*, 4(8), pp.57-66.

HILLA, J.O., AND SUN, M. 1993. A method for measuring mechanical work and work efficiency during human activities. *Journal of Biomechanics*, 26(3), pp.229-41.

HIMBER, J. et al. 2001. Time series segmentation for context recognition in mobile devices. In *First IEEE International Conference on Data Mining (ICDM'01)*, San Jose, CA, 2001. IEEE Computer Society. pp.203-10.

HOSEINITABATABAEI, S.A., GLUHAK, A., AND TAFAZOLLI, R. 2011. uDirect: A novel approach for pervasive observation of user direction with mobile phones. In *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, Seattle, 2011. pp.75-83.

HUYNH, T.M., AND SCHIELE, B. 2008. Discovery of activity patterns using topic models. In Proceedings of *the 10th international conference on Ubiquitous computing*, Seoul, 2008. pp.10-19.

ICHIKAWA, F., CHIPCHASE, J., AND GRIGNANI, R. 2005. Where's the phone? A study of mobile phone location in public spaces. In *International Conference on Mobile Technology, Applications and Systems, 2005 2nd*, Guangzhou, 2005. Nokia. pp.1-8.

ISO, T., AND YMAZAKI, K. 2006. Gait analyzer based on a cell phone with a single three-axis accelerometer. In *Mobile HCI'06*, Helsinki, 2006. ACM. pp.141-44.

JAIN, A.K., AND MURTY, M.N. 1999. Data Clustering : A review. *ACM Computing Surveys*, 31(3), pp.265-323.

KANJO, E. 2010. NoiseSPY: a real-time mobile phone platform for urban noise monitoring and mapping. *Mobile Networks and Applications*, 15(4), pp.562-74.

KANJO, E., BACON, J., AND ROBERTS, D. 2009. MobSens: making smart phones smarter. *Pervasive Computing*, 8(4), pp.51-56.

KANSAL, K., 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)*, Urbana-Champaign, IL, 2007. ACM.

KAPADIA, A., KOTZ, D., AND TRAINDOPOULOS, N. 2009. Opportunistic sensing: security challenges for for new paradigm. In *First International Communication Systems and Networks and Workshops, 2009. COMSNETS 2009.*, Bangalore, 2009. pp.1-10.

KATZAKIS, N., AND HORI, M. 2009. Mobile phone as 3-DOF controllers: A Comparative Study. In *Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing*, Chengdu, 2009. IEEE Computer Society. pp.345-49.

KAWAHARA, Y., KURASAWA, H., AND MORIKAWA, H. 2007. Recognizing user context using mobile handsets with acceleration sensors. In *IEEE International Conference on Portable Information Devices, 2007. PORTABLE07*, Orlando, FL, 2007. IEEE. pp.1-5.

KAWAHARA, Y., KURASAWA, H., AND MORIKAWA, H. 2007. Recognizing user context using mobile handsets with acceleration sensors. In *In (IEEE) Intl. Conf. on Portable Information Devices (PORTABLE'07).*, 2007. IEEE. pp.1-5.

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*, San Diego, CA, 2008. pp.457-66.

KIM, S.J., AND LEE, C.Y. 1996. Modeling and analyzing of the dynamic location registration and paging in microcellular systems. *IEEE Transaction on Vehicular Technology*, 45(1), pp.82-90.

KÖNÖNENA, V. et al. 2010. Automatic feature selection for context recognition in mobile devices. *Pervasive and Mobile Computing*, 6(2), pp.181-97.

KOUROGI, M., AND KURATA, T. 2003. Personal positioning based on walking locomotion analysis with self-contained sensors and a wearable camera. In *Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, Tokyo, 2003. IEEE Computer Society.

KOUROGI, M., AND KURATTA, T. 2003. A wearable augmented reality system with personal positioning based on walking locomotion analysis. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, Tokyo, 2003. IEEE Computer Society. p.342.

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

KUNZE, K., AND LUKOWICZ, P. 2007. Using acceleration signature from everyday activities for on-body device location. In *Proceedings of the 2007 11th IEEE International Symposium on Wearable Computers*, Boston, 2007. pp.1-2.

KUNZE, K., LUKOWICZ, P., JUNKER, H., AND TROSTER, G. 2005. Where am I: Recognizing on-body positions of wearable sensors. In *Location and Context Awareness*. Springer, Springer Berlin / Heidelberg, pp.264-75.

KUNZE, K., LUKOWICZ, P., PARTRIDGE, K., AND BEGOLE, B. 2009. Which way am i facing: Inferring horizontal device orientation from an accelerometer signal. In *Wearable Computers, 2009. ISWC '09. International Symposium on*, Linz, 2009. pp.149-50.

KWAPISZ, J.R., WEISS, G.M., AND MOORE, S.A. 2010. Activity recognition using cell phone accelerometers. In *SensorKDD'10*, Washington, DC, 2010. ACM.

KYRIAZAKOS, S.A., AND KARETSOS, G.T. 2000. Architectures for the provision of position location services in cellular networking environments. In *Telecommunications and IT Convergence Towards Service E-volution*. J. Delgado et al., Eds. Springer Berlin / Heidelberg,

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

LAFORTUNE, M.A. 1991. Three dimensional acceleration of tibia during walking and running. *Journal of Biomechanics*, 24, pp.877-86.

LANE, N.D. et al. 2010. A survey of mobile phone sensing. *IEEE Communication Magazine*, 48(9), pp.140-50.

LARSEN, J.E., AND LUNIEWSKI, M. 2009. Using mobile phone contextual information to facilitate managing image collections. In *Personal Information Management*, Vancouver, BC, 2009.

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

LEE, S.W., AND MASE, K. 2002. Activity and Location Recognition Using Wearable Sensors. *IEEE Pervasive Computing*, 1, pp.24-32.

LERCH, A., 2009. *Software-based extraction of objective parameters from music performances*. PhD Thesis. Berlin: Technical University Berlin.

LESTER, J., CHOUDHURY, T., AND BORRIELLO, G. 2006. A practical approach to recognizing physical activities. In *Lecture Notes in Computer Science: Pervasive Computing*. K. Fishkin, B. Schiele, N. Nixon & A. Quigley, Eds. Springer Berlin / Heidelberg,

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*, Beijing, 2007. pp.100-03.

LIU, Q., MCEVOY, P., AND LAI, C.J. 2006. Mobile camera supported document redirection. In *Proceedings of the 14th ACM International Conference on Multimedia,* Santa Barbara, CA, 2006. pp.791-92.

LOCKHEED, M.C., AND NASHUA, N. 1996. Real-time discrimination of broadcast speech/music. In *IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP-96*, Atlanta, 1996. pp.7-10.

LOTTE, F. et al. 2007. A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 7(3), pp.1-13.

LUINGE, H.J., VELTINK, P.H., AND BATEN, C.T.M. 1999. Estimation of orientation with gyroscope and accelerometers. In *Proceedings of The First Joint BMES/EMBS Conference Saving Humanity, Advancing Technology*, Atlanta, 1999. IEEE. pp.884-85.

LU, H. et al. 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*, Karkow, 2009. pp.165-78.

LU, H. et al. 2010. The Jigsaw continuous sensing engine for mobile phone applications. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, New York, NY, 2010. pp.71-84.

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

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

MATHIE, M.J., COSTER, A.C.F., LOVELL, N.H., AND SELLER, B.G. 2004. Accelerometery: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiological Measurement*, 25, pp.1-20.

MAURER, U., ROWE, A., SMAILAGIC, A., AND SIEWIOREK.D 2006. Location and activity recognition using eWatch: a wearable sensor platform. In *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg,

MAYAGOITIA, R.E., LOTTERS, J.C., VELTINK, P.H., AND HERMENS, H. 2002. Standing balance evaluation using a triaxial accelerometer. *Journal of Gait and Posture*, 16, pp.55-59.

MCNIELL, 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 Technical Journal*, 12(2), pp.79-86.

MILUZZO, E. et al. 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*, Raleigh, 2008. ACM. pp.337-50.

MIRKIN, B. 2005. *Clustering for Data Mining, a data recovery approach*. Chapman & Hall/CRC, Boca Raton.

MISHRA, A.R. 2004. *Fundamentals of cellular network planning and optimization*. John Wiley & Sons, Chichester.

MIZELL, D. 2003. Using gravity to estimate accelerometer orientation. In *IEEE International Symposium on Wearable Computers*, Washington, DC, 2003. IEEE Computer Society. p.252.

MOHAN, P., PADMANABHAN, V.N., AND RAMJEE, R. 2008. Nericell: rich monitoring of road and traffic conditions using mobile smartphones. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, Raleigh, 2008. pp.323-36.

MORRIS, S., AND PRADISO, J. 2002. Shoe-integrated sensor system for wireless gait analysis and real-time feedback. In *Proceedings of the 2nd Joint IEEE EMBS and BMES Conferences*, Houston, 2002. pp.2468-69.

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 2004*, 2004. pp.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*, Cologne, 2005. ACM. pp.243 - 257.

MUN, M. et al. 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*, Krak, 2009. pp.55-68.

MURRAY, M. 1967. "Gait as total pattern of movement". *American journal of physical medicine*, 46(1), pp.290-333.

NOKIA 2005. Workshop on large-scale sensor networks and applications. Kussamo, 2005. Nokia. pp.3-6.

OFSTAD, A., NICHOLAS, E., SZCODRONSKI, R., AND CHOUDHURY, R.R. 2008. AAMPL: accelerometer augmented mobile phone localization. In *International Conference on Mobile Computing and Networking*, California, 2008. ACM. pp.13-18.

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,* Bangkok, 2005. pp.1104 - 1115.

OLGUIN, D.O., AND PENTLAND, A. 2006. Human activity recognition: accuracy across common locations for wearable sensors. In *IEEE 10th symposium on Wearable Computers*, Montreaux, 2006.

OLGUIN, D.O., AND PENTLAND, A. 2008. Social sensors for automatic data collection. In *14th Americas Conference on Information Systems.*, Toronto, 2008. pp.1-10.

OLMEDILLA, D., MARTÍNEZ, E.F., AND LARA, R. 2010. Mobile web profiling: a study of off-portal surfing habits of mobile users. In *Lecture Notes in Computer Science:USER MODELING, ADAPTATION, AND PERSONALIZATION*. P.D. Bra, A. Kobsa & D. Chin, Eds. Springer Berlin / Heidelberg,

ONNELA, J.P. et al. 2007. Structure and tie strengths in mobile communication networks. In *Proceedings of the National Academy of Science of the United States of America (PNAS)*, 2007. pp.7332-36.

PAULOS, E., AND GOODMAN, E. 2004. The familiar stranger: anxiety, comfort, and play in public places. In *Conference on Human Factors in Computing Systems*, Vienna, 2004. ACM, New York USA. pp.223-30.

PEEBLES, D. et al. 2010. Community-guided learning: exploiting mobile sensor users to model human behavior. In *24th National Conference in Artificial Intelligence*, Atlanta, 2010.

PENTLAND, A. 2009. *Honest signals : how they shape our world*. Massachusetts Institute of Technology,

PICARD, R.W., AND HEALEY, J. 1997. Affective wearables. In *First International Symposium on Wearable Computers, 1997. Digest of Papers.*.,

PIETQUIN, O., n.d. *A framework for unsupervised learning of dialogue strategies*. ISBN: 2-930344-63-6 PHD Thesis. Universitaires de Lovain.

PREECE, S.J. et al. 2009. Activity identification using body-mounted sensors - a review of classification techniques. *Physiological Measurement*, 30, pp.1-33.

JUDD, T., AND LEVI, R., 1996. *Dead reckoning navigational system using accelerometer to measure foot impacts*. US Patent Number 5,583,776.

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*, 2005.

RAVI, N. et al. 2005. Indoor localization using camera phones. In *Proceedings of the Seventh IEEE Workshop on Mobile Computing Systems and Applications*, Orcas Island, WA , 2005. p.19.

REDMOND, D.P., AND HEGGE, F.W. 1985. Observation on the design and specification of a wrist-worn human activity monitoring system. *Behavior Res. Methods, Instruments & Computer*, 17, pp.659-69.

ROSS, T. 2004. *Fuzzy Logic with Engineering Applications*. John Wiley & Sons,

RUF, B., AND DETYNIECKI, M. 2009. Identifying paintings in museum galleries using camera mobile phones. In *Proceedings of the Singaporean-French IPAL Symposium*, 2009. pp.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 *Sixth International Conference on Networked Sensing Systems (INSS)*, Pittsburgh, 2009. pp.1-9.

SANTOS, A.C. et al. 2010. Providing user context for mobile and social networking applications. *Pervasive and Mobile Computing*, 6(3), pp.324-41.

SANTOS, A.C. et al. 2009. Context inference for mobile applications in the UPCASE project. In *Mobile wireless middleware, operating systems, and applications*. Springer Berlin Heidelberg,

SASHIMA, A. et al. 2008. CONSORTS-S: A mobile sensing platform for context-aware services. In *International conference on intelligent sensors (ISSNIP) 2008*, Sydney, 2008. pp.417-22.

SCHEIRER, E., AND SLANEY, M. 1997. Construction and evaluation of a robust multifeature speech/music discriminator. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-97*, Munich, 1997. pp.1331-34.

SCHILLER, J. 2003. *Mobile Communications*. PEARSON, Harlow.

SEKINE, M. et al. 2002. Discrimination of walking patterns using wavelet based fractal analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 10(3), pp.188-96.

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*, Chicago, IL, 2000. pp.1356-59.

SIEWIOREK, D. et al. 2003. SenSay: a context-aware mobile phone. In *Proceedings of the 7th IEEE International Symposium on Wearable Computers.*, 2003. IEEE Computer Society Washington. p.248.

SOHN, T. et al. 2005. A study of location-based reminders on mobile phones. In *Seventh International Conference on Ubiquitous Computing Ubicomp'05*, Tokyo, 2005. pp.232-50.

SOHN, T. et al. 2006. Mobility detection using everyday GSM traces. In *Proceedings of the Eighth International Conference on Ubiquitous Computing*, Orange County, 2006. pp.212-24.

STIEFMEIER, T., OGRIS, G., AND LUKOWICZ, P. 2008. Wearable activity tracking in car manufacturing. *Pervasive Computing*, pp.42-50.

SUH, Y., SHIN, C., AND WOO, W. 2009. A mobile phone guide: Spatial, personal, and social experience for cultural heritage. *Consumer Electronics, IEEE Transactions on*, 55(4), pp.2356-64.

TAKACS, G. et al. 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*, Vancouver, 2008. pp.427-34.

TITTERTON, D.H., AND WESTON, J.L. 2002. *Strapdown inertial navigation technology*. Institution of Electrical Engineers,

VAHDATPOUR, A., AMINI, N., AND SARRAFZADEH, M. 2009. Toward unsupervised activity discovery using multi-dimensional motif detection in time series. In *Proceedings of the 21st international joint conference on Artificial intelligence*, San Francisco, 2009. pp.1261-66.

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

VIEIRA, M.R. et al. 2010. Querying spatio-temporal patterns in mobile phone-call databases. In *Proceedings of the Eleventh International Conference on Mobile Data Management*, Kansas City, MO, 2010. IEEE Computer Society. pp.239-48.

VINICIARELLI, A., PANTIC, M., AND BOURLAND, H. 2009. Social signal processing : Survey of an emerging domain. *Image and Vision Computing*, 27(12), pp.1743-59.

WAGNER, D. et al. 2010. Real-time detection and tracking for augmented reality on mobile phones. *IEEE Transactions on Visualization and Computer Graphics*, 16(3), pp.355-68.

WANG, Y. et al. 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*, Karakow, 2009. pp.179-92.

WANG, J., ZHAI, S., AND CANNY, J. 2006. Camera phone based motion sensing : Interaction techniques, applications and performance study. In *ACM Symposium on User Interface Software and Technology*, Montreux, 2006. pp.101-10.

WEBB, A. 1999. *Statistical pattern recognition*. Wiley, Chichester.

WITTKE, M. et al. 2009. Activity recognition using optical sensors on mobile phones. In *GI Jahrestagung* , Lübeck, 2009. pp.2181-94.

WOODMAN, O.J., 2007. *An introduction to inertial navigation*. ISSN: 1476-2986 Technical Report. Cambridge: University of Cambridge.

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

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*, Beijing, 2009. ACM. pp.1-10.

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

YI, J.S., CHOI, Y.S., AND JACKO, J.A. 2005. Context awareness via a single device-attached accelerometer during mobile computing. In *Mobile HCI '05*, Salzburg, 2005. ACM. pp.303-06.

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

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

YONG-XIANG, Z., WEI-GONG, Z., XIAO-XU, Z., AND HUI-MEI, Y. 2009. Study on electronic image stabilization system based on MEMS gyro. In *International Conference on Electronic Computer Technology*, Macau, 2009. pp.641-43.

ZHANG, S., YUAN, C., AND ZHANG, Y. 2008. Handwritten character recognition using orentation quantization based on 3D accelerometer. *MobiQuitous*, pp.21-25.

ZHAO, Y. 2000. Mobile Phone Location Determination and Its Impact on Intelligent Transportation Systems. *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, 1, pp.55-64.

ZINNEN, A., BLANKE, U., AND SCHIELE, B. 2009. An Analysis of Sensor-Oriented vs. Model-Based Activity Recognition. In *International Symposium on Wearable Computers, 2009. ISWC '09*, Linz , 2009. pp.93 - 100.

ZINNEN, A., WOJEK, C., AND SCHIELE, B. 2009. Multi activity recognition based on body model-derived primitives. In *Location and Context Awareness*. Springer Berlin / Heidelberg,

ZULKEFLY, S.N., AND BAHARUDIN, R. 2009. Mobile phone use amongst students in a university in malaysia: Its correlates and relationship to psychological health. *European Journal of Scientific Research*, 37(2), pp.206-18.

APPENDIX.1

Table.2. Comparison of several systems prototypes for user context recognition

System	Sensor(s)	Sensor node	Processing Units	Sampling rate(Hz)	Pre-processing	Context inference	Context	Accuracy	Goal
SurroundSense [Azizyan et al., 2009]	Camera, Microphone, Accelerometer, Wi-Fi	Nokia N95	Nokia N95	Accelerometer (<4), Camera(.2), Microphone 8Khz, Wi-Fi (.2)	Normalization, Average, Variance, HSL, Colour, 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 + back-end server	Using power aware 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-mean clustering, (on Smartphone)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., 2010]	Accelerometer	Smartphone	Back end 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.	Multi-Modal sensor board(MSB)	Back end server	4Hz	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.
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 Energy efficient sensing system for mobile phones.

System	Sensor(s)	Sensor node	Processing Units	Sampling rate(Hz)	Pre-processing	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, Uninterruptible, Active and default.	-	Provides a context aware 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, Bays rule, GMM	Location pattern, Proximity pattern	(95 %)Identify next location, 90% (face to face contacts), 90% (relationships)	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 + back-end server	Once every 5 min	Updating thresholds and weights sent by user o 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, 2006]	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)	Back end 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
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 a cartoon based information collected via mobile devices like Smartphone.

System	Sensor(s)	Sensor node	Processing Units	Sampling rate(Hz)	Pre-processing	Context inference	Context	Accuracy	Goal
[Gyorbiro 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	Back end 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, Wi-Fi	External sensors	PDA(Sharp Zaurus SL6000L)	Accelerometer (90), Microphone(8), Wi-Fi(.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
Soundsense [Lu et al., 2009]	Microphone	Apple Iphone	Smartphone	8000	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	Markove model, Decision tree (J48) Gaussian discriminative model	Human Voice, music, ambient	>78%	Recognizing everyday life sound events on mobile phone.