UNIVERSITÀ DI PISA

Scuola di Ingegneria

Corso di Laurea in Ingegneria delle Telecomunicazioni

Master Thesis

An Energy-Efficient Framework for Data Collection Optimization in Mobile Crowd Sensing Systems

Advisor
Prof. Stefano Giordano

External Supervisors
Dr. Dzmitry Kliazovich
Claudio Fiandrino

Candidate
Andrea Capponi

ANNO ACCADEMICO 2014-2015

Summary

Mobile Crowd Sensing (MCS) has recently gained popularity becoming an appealing paradigm for sensing and collecting data. It is an emerging area of interest for researchers as smartphones are becoming ubiquitous devices in use around the world. The key fact is that mobile devices have not only computing and communication resources, but they also offer the possibility to exploit a rich set of sensors for enabling new applications across a large variety of domains.

Considering the latest generation sensors available in modern mobile devices such as smartphones, wearable and IoT devices, this thesis proposes a taxonomy to analyze and categorize sensors according to their implementation scope, sampling activity and purpose.

In MCS systems users contribute data gathered from sensors embedded in mobile devices, including smartphones, tablets and wearable devices. Consequently, it is required participation and contribution of a large number of users to guarantee the efficiency of the system. Mobility and intelligence of human participants guarantee higher coverage and better context awareness, compared to traditional sensor networks. On the other hand, individuals may be reluctant to share data and it is needed to introduce incentives. This work analyzes in depth different approaches of the sensing process and elaborate a novel detailed taxonomy that provides a clear organization of the state-of-the-art. The layered taxonomy includes definitions and comprehensive analysis of the available MCS solutions, providing a hierarchical view. On the first level it classifies the works according to their application target. On the second level, the taxonomy defines several categories that capture in detail the properties of MCS systems and analyzes the works in relation to their application target. Later, a classification overviews MCS systems defining existing methodologies, architecture design and applications organized according to the introduced taxonomy.

Several research works focus on sensing applications or incentive mechanisms, while data collection requires a more detailed investigation. In this thesis a novel distributed algorithm is proposed for gathering information in cloud-based mobile crowd sensing systems. The objective is to minimize the cost of sensing and reporting processes and in the meanwhile to maximize the utility of data

collection, for instance the quality of contributed information. The performance of the proposed framework is evaluated analytically considering different metrics, such as the average number of samples collected per area.

The last part of the thesis consists in the development of a simulator for MCS systems. The simulator is a discrete-event simulator in which the participants contribute data to the MCS system opportunistically, that is they collect data without being an active part in the sampling process. In the simulator, the participants move in a real city environment (Luxembourg City center was chosen for the purpose) with pedestrian mobility and generate data by means of latest generation sensors commonly available in today smartphones.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Nowadays smartphones are ubiquitous mobile devices, embedded with a set of powerful sensors that could be employed for many applications. In the last years, number of smartphones in use worldwide broke the 1 billion mark [1] and the total shipments showed a strong growth, for instance in 2014 up 28.4 percent from 2013 [2]. In 2015 the sales of smartphones to end users reached 1.4 billion units, up 9.8 percent from 2014 [3] and according to a new mobile phone forecast from the International Data Corporation (IDC) Worldwide Quarterly Mobile Phone Tracker, smartphone shipments are expected to grow 7.4 percent in 2016. Furthermore, IDC predicts that smartphones sales will continue to grow, even if an ever decreasing rate and the worldwide shipment volumes are forecast to reach 1.9 billion units annually by 2019 [4].

Sensing is the enabling factor for developing applications across a large variety of domains, such as home care, health care, social networks, public safety, environmental monitoring and intelligent transportation systems. Fixed sensors provide information for specific areas of interest (e.g., video surveillance) and have problems in management and maintenance. On the contrary, mobile devices provide unlimited possibilities. First, they have a rich set of sensors, considering both the ones that are built-in and the possibility to connect others to the smartphone. Second, users move by themselves, have interest in the maintenance of mobile devices and recharge them. Consequently, smartphones are essential for developing applications to help and change life of the people (e.g., in public safety for car accidents, crime detection or natural disasters). For this reason, the analysis will focus on specialized sensors such as ambient light sensor, accelerometer, digital compass, gyroscope, GPS, proximity sensor and general purpose sensors like microphone and camera.

Taking into account that smartphones are not only personal sensing platforms, but also computing and communication devices, the concept of Mobile Crowd Sensing (MCS) will be introduced. It is a new sensing paradigm in which mobile users can easily and efficiently collect and share sensing information, in order to enable numerous distributed and large scale applications. MCS can greatly improve citizens everyday life and provide new perspectives to urban societies. MCS is an essential solution for building smart cities of the future, which aim at using ICT solutions to improve management of everyday life of their citizens [5, 6]. Several MCS-based solutions have already been proposed in this sense. For instance, the accelerometer and the GPS may be exploited for road maintenance [7] or the GPS and the camera for checking free spots in a car park [8, 9]. MCS is also referred as people- or human-centric sensing paradigm because human involvement is the most important feature in the process. Indeed, system devices are no longer owned and managed by a single authority but belong to people with different interests who live and move in several contexts. Consequently, sensing data is more related to interactions between people and their surroundings rather than some physical phenomena of interest to be monitored. Furthermore, system devices have much more powerful resources and maintenance than sensor nodes and people typically charge them on daily basis. The involvement of people, who are not just passive users but also active data contributors, brings not only huge advantages but also points out many problems, which will be analyzed and some possible solutions will be proposed.

Energy efficiency is another fundamental aspect to take into account, because the primary usage of smartphones should be reserved for the users' regular activities. For this reason, it is fundamental that sensing applications would not introduce significant energy consumption. Contributing data should not consume too much battery and prevent the users from accessing their usual services, such as instant messaging. For instance, the Global Positiong System (GPS) consumes a significant amount of energy, much higher than other sensors, such as the accelerometer or the gyroscope. Consequently, it is fundamental to make use of these sensors efficiently. One of the most challenging problem to overcome is the position of the device, indeed sensors will be placed in a way most convenient to the user (e.g., in a pocket) and not necessarily in a manner most conductive to high fidelity data gathering for an application. In these cases, sensing could be only a waste of energy, so it is necessary to have an idea of the environmental context of the device. The aim is to maximize the utility of the sensors' usage and minimize the waste and the cost, in terms of energy and data overhead.

1.2 Contributions

The proposal of this thesis can be summarized as follows: it analyzes sensing elements, MCS systems and architectures, introduces a novel hierarchical taxonomy for MCS systems, develops a distributed framework for data collection and presents a discrete-event simulator in a realistic city environment.

To easy and simplify the understanding of MCS, this thesis elaborates a novel taxonomy that provides a clear organization of the state-of-the-art. The proposed taxonomy, which is layered, gives a hierarchical view and exploits as the key concept to classify existing studies the purpose of the application. Furthermore, it defines several categories to capture in detail properties of MCS systems and analyze works in relation to their application target.

A new framework is introduced and the process of data collection is analyzed from a new point of view, defining an algorithm to maximize the utility of data collection in MCS and to minimize the costs. Both the smartphone and the collector side are considered. The aim is to decide when to sample and report the data generated by smartphone sensors, taking into account different parameters from both sides. In other words, the algorithm introduces a matching process for taking the decision of sampling and reporting. Performance are evaluated analytically taking into account different metrics.

The last objective of the thesis is to assess the efficiency of the framework taking into account a large number of participants contributing data in a real city. For the purpose, a custom simulator called CrowdSenSim is developed. Crowdsensim is a discrete-event simulator that supports pedestrian mobility. The participants move in a realistic city environment and communicate with base stations, which are in the real position given by latitude and longitude.

1.3 Thesis Organization

This thesis is organized as follows:

- Chapter 2 provides an overview of smartphone sensors, from the general purpose to the more specific, defining which ones are always on and consequently more conservative and the others that consume more energy. It presents a distinction between embedded and non-embedded sensors: the first, which can be found more or less in each device, are limited and well-known, while the second, which can be connected to the device via bluetooth, are several and more oriented to the single user choice, consequently to commerce. Furthermore, the chapter introduces the concept of context as a central role in our dissertation, taking into account smartphone position (e.g., pocket), user mobility and utilization of the device. The chapter ends presenting some applications, which exploit the interaction between sensors.

- Chapter 3 describes a new sensing paradigm, which is called Mobile Crowd Sensing and highlights its importance for developing more and more useful and important application in everyday life. It presents a little survey, taking into account the main features of this new paradigm and the differences of approaches in presented models and applications. For instance, it focuses on the differences between a participatory or an opportunistic approach, on the assigned tasks, or also on which incentives can be used to make the users collaborative as volunteers.

- Chapter 4 presents a new energy-efficient framework for data collection optimization in mobile crowd sensing systems. The aim is to devise a framework that minimizes energy-consumption and maximize utility of collected data. To achieve this goal, it is presented a framework that takes into account several parameters, including data collection utility, smartphone sensing potential, environmental context and battery level. This section also discusses possible enhancements of the current model.

- Chapter 5 evaluates the performance of the proposed algorithm analytically. It is explained the configuration of the tools used and some metrics are presented to simulate different situation in details. The results are discussed and analyzed to understand the achievements of the model in the presented scenario.

- Chapter 6 introduces CrowdSenSim, a new simulator that is developed in a realistic city environment. It relies on simulating the movement of pedestrians who sense and report data in a real city. Starting from the real position of LTE base stations and WiFi hotspots in the city of Luxembourg and exploiting some models of users' mobility all around the city, the objective of the simulator consists in measuring the costs of the devices experience and determining the amount of contributed data.

- Chapter 7 discusses about the conclusion and presents some ideas about future works.

Chapter 2

Sensing Elements

Sensing is the enabling factor for developing applications across a large variety of domain, such as home care, health care, social networks, safety, noise mapping, environmental monitoring and intelligent transportation systems. For this reason, it is fundamental to analyze sensors, which are at the heart of any sensing systems, including MCS systems. Unlike for the traditional case with fixed sensors, in MCS systems the devices are no longer owned and managed by a single authority. Participants have usually different interests and they can decide independently if contributing or not in the process. As a consequence, sensing is a process concerning the contextual activities that happen in the surrounding of the participants. In traditional fixed sensor systems, sensing focus instead on a particular phenomena of interest that needs to be monitored. In other words, people are no longer just passive data users, but also active data contributors. Furthermore, due to the fact that users take care of their smartphones, system devices have much more powerful resources and maintenance than typical sensor nodes and can be charged regularly. Accelerometer, gyroscope, GPS, microphone and camera are just a few examples of typical built-in sensors in modern smartphones, wearables and IoT devices. Sensors are responsible of gathering data, which is often delivered to a collector or used for local application. Analyzing these sensing elements there are some challenges to face with. First of all, relying on a people-powered mobile architecture means that the sensing devices characteristics are heterogeneous, because different sensor types are embedded in an infinite number of devices that have several storage, processing and communication capabilities. Moreover, sensors will be placed in a way most convenient to the user (e.g., in a pocket) and not necessarily in a manner most conductive to high fidelity data gathering for an application. This section first illustrates a taxonomy on sensors and then analyzes and classifies sensing elements according to the proposed taxonomy.

2.1 Taxonomy on Sensing Elements

Sensing elements can be broadly categorized into three categories according to their implementation scope, continuity of sensing activity and type of sensing information. Fig. 2.1 illustrates this classification. *Implementation* category distinguishes between sensors embedded or not in the mobile devices. The *sampling activity* differentiates between sensors that work continuously because they are necessary for basic operations of the device and those that require user intervention to become active. Finally, the *purpose* category investigates the applications the sensors are designed for. This classification unveils a number of properties. For example, passive sensors perform sensing continuously and are not manually activated by the user such as the camera. For this reason, they are typically embedded in the mobile devices and they operate consuming a small amount of energy. Having such deep understanding helps devising applications using sensor resources properly, such as exploiting a passive sensor to switch on or off another one. For instance, turning off the screen using the proximity sensor helps saving battery lifetime as the screen is a major cause of energy consumption [10]. Turning off the ambient light sensor and the GPS when the user is not moving or indoor, allows to obtain additional power savings.

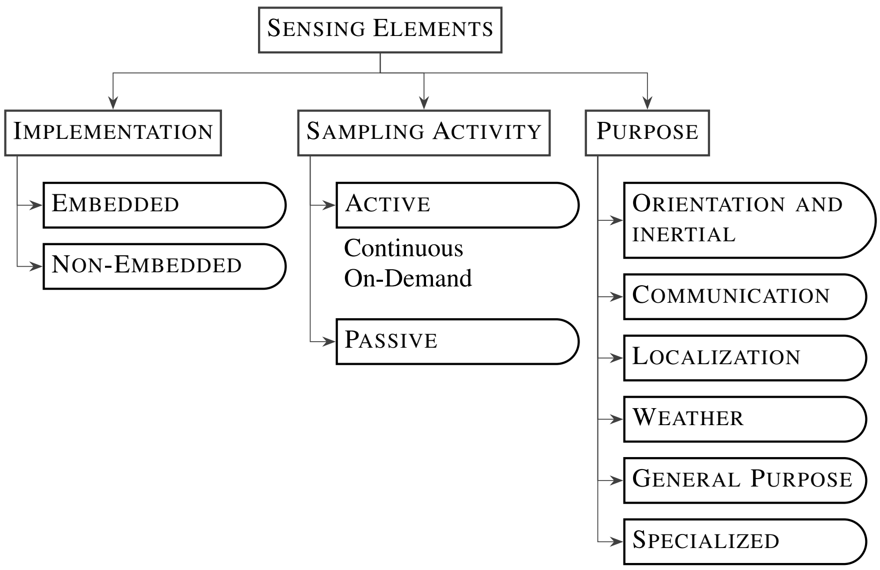


Figure 2.1: Taxonomy on mobile devices sensors

Implementation Scope

The vast majority of available sensors are embedded in mobile devices. Nevertheless, non-embedded sensing elements exist. Indeed, they are designed for very specific purposes and vendors have no interest in large scale production. For instance, the gluten sensor comes as a standalone device and it is designed to work in couple with smartphones, which are responsible to receive, store and process food records of gluten detection [11]. These tasks can only be accomplished using wireless connection. Therefore it becomes necessary to distinguish between sensors that are embedded in the devices and sensors that are standalone devices.

Embedded Sensors

Integrating sensors into mobile devices is nowadays common practice. If embedded, sensors do not require to be paired with other devices for data delivery. Sensors are essential for ordinary operations of smartphones (e.g., the microphone for phone calls), for social purposes (e.g., the camera to take pictures and record videos) and user applications (e.g., GPS for navigation systems). Typically, smartphones are equipped with a higher number of embedded sensors than wearable devices.

Non Embedded Sensors

These sensors are typically standalone devices and are designed to be paired with a smartphone for data transmission. Indeed, they are very small devices with limited storage capabilities. Communications rely on wireless technologies like Bluetooth or Near Field Communications (NFC). Nevertheless, particular sensors such as the GasMobile hardware architecture can be wired connected with smartphones through USBs [12]. Whereas embedded sensors are used only for popular applications, the design of non-embedded sensor is very specific and which either single individuals or the entire community can profit. To illustrate, only the celiac community can take advantage of the gluten sensor [11] while an entire city can benefit from the fine dust sensors [13] or nuclear radiation monitoring [14].

Sampling Activity

Nowadays a growing number of application people use on daily basis require sensors to operate. Moreover, mobile devices themselves use sensors for basic functionalities. For example, auto-adjusting the brightness of the screen require the ambient light sensor being active, understanding the orientation of the device can be possible only having accelerometer and gyroscope working continuously. On the other hand, a number of sensors can be switched on and off manually by user intervention: taking a picture or recording a video requires the camera being active only for a while. It should be noted that some always-on sensors can be used temporarily by user application such as games.

The second classification allows to categorize sensors on the basis of their continuity in sampling activity.

Passive Sensors

Passive sensors are required to accomplish mobile devices basic functionalities, such as detection of rotation and acceleration. As a consequence, they are typically embedded sensors, they run continuously and consume a very little amount of energy. For such a reason, it has become convenient for a number of applications to make use of these sensors in other context. For instance, the accelerometer can be employed for user-activity recognition [15, 16] and to monitor the driving style [17] if used in pair with the gyroscope. Furthermore, it is possible to use passive sensors for context-awareness detection, such sensing user surroundings [18].

Active Sensors

Active sensor require user intervention to become active and typically serve more complex application functionalities than passive sensors. As a consequence, they consume much more energy and for this reason they are typically disabled for power savings. The GPS, the camera and the microphone are representative examples. These sensors can be attributed to the subcategory *continuous* as, once they are active, they provide readings at a given sampling rate until the user switch them off. For instance, the GPS when it is used in cooperation with a navigation system, the camera and the microphone for recording videos and audio respectively. Moreover, they can be attributed to the subcategory *on-demand* because when they provide one reading only, after being active. Typical examples are the camera for taking picture, the microphone to reveal the level of noise in dB and the GPS to sense the exact position of a mobile device while sensing something (e.g., petrol prices in a gas station located anywhere [19]).

Purpose

Classifying sensors according to their purpose is fundamental to understand in which applications they can serve. Very often applications require multi-sensing capabilities. To illustrate, DietSense [20] captures pictures of food and reports thanks to the GPS location and timestamp references of the sample. In addition, the microphone detects in which environment the sample was collected. The GPS can be attributed to a *localization* category: it allows to detect the position of the device. The same functionality can be achieved using WiFi. This is especially true for indoor environments, where the GPS is not working. The camera and the microphone are *general purpose* sensors. Although they provide a specific functionality, image/video and audio recording respectively, the objective while performing sensing can vary. Recalling the DietSense and the GasMobile examples, images are employed for both healthcare and price detection. In contrast, *specialized* sensors allows gathering information useful for one purpose only. The gluten belongs to this category.

Orientation and Inertial

Inertial sensors measure acceleration and rotation, expressed as vector values. These sensors are the accelerometer and the gyroscope. Sometimes, also the magnetometer, which measures magnetic fields, is considered to be an inertial sensors. Magnetometer, accelerometer and gyroscope form the Inertial Measurement Unit (IMU). To illustrate, all these sensors need to be aligned for obtaining accurate 3D orientation estimates of the device [21].

Communication-related

In first place, WiFi, Bluetooth and NFC enable connectivity of mobile devices to the Internet or with other mobile devices including non-embedded sensors. These sensors can also be employed for locating users nearby [22] and acquiring device location [8].

Weather

Sensors belonging to this category allow to obtain readings useful for environmental monitoring like temperature, pressure and humidity. Among the three types of sensors, the temperature one is the most common implemented in mobile devices.

Localization

GPS, WiFi and Cellular Tower Signal (CTS) are employed to localize mobile devices. The most accurate sensor is the GPS, which consumes a considerable amount of energy when active [23, 10]. Readings obtained from WiFi and CTS are less precise in comparison with GPS [23].

General Purpose

This category groups sensors that can serve for multiple purposes or that are engaged in typical phone operations and can also be employed for sensing. To illustrate, the microphone is essential for phone calls. In addition, it helps in monitoring noise [24] and in context detection [20]. Other sensors belonging to this category are the camera, the ambient light and the proximity sensors.

Specialized

Specialized sensors are application-oriented and are used only in a specific field of potential function, for instance air pollution, allergy food detection or radiation monitoring. Typically specialized sensors are non-embedded. Being targeted to a particular application, the market demand is limited and therefore vendors have little interest in implementing them for large distribution. Moreover, if embedded they have a potential to drain the battery very quickly. Examples of specialized sensors are the gluten sensor [11], the fine dust sensor [13] and the radiation sensor [14].

2.2 Classification

This section groups and classifies the most common and available on the market mobile devices sensors. Table 2.1 illustrates and summarizes their properties according to the taxonomy illustrated in Subsection 2.1.

Accelerometer

The accelerometer measures the non-gravitational acceleration of the device. It is a dynamic and inertial sensor for motion detection, which defines the position of the device in the space through x, y and z axis. Being fundamental for proper operation of the devices, the accelerometer is very cheap in terms of energy consumption and embedded in all mobile devices, including IoT devices. It is a passive sensor which generates samples at low frequencies, typically in the order of 40 Hz. In smartphones, the accelerometer makes possible screen auto-rotation and trigger context-based operations, i.e. turning off the GPS when the user is not moving. Moreover, accelerometers contain highly detailed information about phone movement, enabling fine-grained distinction of different activity or transportation modalities. Activity recognition such as detection of movement patterns (e.g. walking/running [15, 25, 26]) or actions (e.g. driving, riding a car or sitting [27, 28, 29]) is a very important feature that the accelerometers enable. For instance, it can be employed for health care to detect a fall and user reaction after a fall [16, 30, 31, 32] and to distinguish transportation modes [33, 34, 35, 36, 37, 38].

Accelerometers measure the proper acceleration, which is the physical acceleration experienced by an object or a person, thus the acceleration relative to a free-fall, or inertial, observer who is temporarily not moving relative to the object being measured. Such accelerations are measured in terms of g-force. Accelerometer typically has different basic specifications, but the most important are:

- Sensitivity: it is the ratio of an electrical output to the mechanical input. It is usually expressed in terms of volts per unit of acceleration under the specified conditions. It is an indicator of the amount of change in output signal for a given change in acceleration. An accelerometer that can be defined as sensitive will be more precise and accurate.
- Frequency response: is the output signal over a range of frequencies where the sensor should be operating. It is specified with respect to a reference frequency that is where the sensitivity is specified.
- Dynamic range: this is the range between the smallest acceleration detectable by our sensor to the biggest. Measures that do not fall in this range are clipped.
- Bandwidth: it is usually measured in Hertz and indicates the limit of the near-unity frequency response of the sensor, or how often a reliable reading can be obtained.

In order to understand the inertial principle of the accelerometer, it is useful to image a box in shape of a cube with walls that are pressure sensitive and a ball inside of it. If this box is considered in a place with no gravitational fields, the ball will float in the middle of the box, but as soon as the box suddenly moves to one direction, the ball will hit the wall to the opposite direction. The measurement of the pressure force gives the value to output. The accelerometer will actually detect a force that is directed in the opposite direction from the acceleration vector and this is why it is called inertial force. In real life, the gravity force must be subtracted before any measurement because the ball will fall on the ground because of the g-force. An accelerometer at rest relative to Earth surface will indicate approximately 1 g upwards, so this must be subtracted and corrections should be made because of the effects caused by Earth rotation relative to the inertial frame. However, the gravity force can be taken as an advantage of detecting the rotation of a device. For instance, when a user rotates his smartphone, the content will switch between portrait and landscape.

Most smartphones typically make use of Micro-Electro-Mechanical Sensors (MEMS) and the three-axis model; they trade large value range for high precision, for instance iPhone4 has range ±2g and precision 0.018g.

In theory, the displacement can be calculated as:

$$d(t) = d_0 + v_0 \cdot t + \iint_0^T a(t)\, dt\, d\tau,$$

where d(t) displacement, d_0 initial displacement, v_0 initial velocity and a(t) acceleration. This equation is a continuous function, but in real world the a(t) is discrete due to sampling. To calculate the displacement according to discrete values:

$$\int_{t(0)}^{t(n)} a(t)\,dt = \sum_{i=1}^{n} \frac{((a(i-1)+a(i))}{2} \cdot \Delta t,$$

where a(i) is the i-th sample and Δt time increment. Velocity and displacement can be calculated as the following:

$$v(i) = v(i-1) + \frac{a(i-1)+a(i)}{2} \cdot \Delta t,$$

$$d(i) = d(i-1) + \frac{v(i-1)+v(i)}{2} \cdot \Delta t.$$

The accelerometer returns a 3-axis value, so a(t) can be calculated as:

$$a(t) = \vec{a_x} + \vec{a_y} + \vec{a_z},$$

where they are vectors.

While the accelerometer is accurate at measuring the displacement of an object, it is not so good to measure the spin movement of the device. Consequently, a gyroscope is often needed and it is defined a "keeper of direction".

Gyroscope

The gyroscope determines the orientation of a device, measuring the rotation rate around an axis. Similarly to the accelerometer, the gyroscope is an embedded and passive sensor (see Table 2.1). When the device is in horizontal position, the gyroscope provides readings whose value is equal to zero. When used in combination with the accelerometer, the gyroscope helps to detect and recognize user movements. In addition, it can be employed for recognition of physical activity [39], driving style [17] or providing a car navigation system [40]. Another very important application based on gyroscope readings is speech recognition [41].

Gyroscope is based on the fundamental principle of the conservation of angular momentum: in any system of particles, the total angular momentum of the system relative to any point fixed in space remains constant, provided no external forces act on the system. It detects the current orientation of the device, or changes in the orientation; more precisely, the orientation can be computed from the angular rate that is detected by the gyroscope, expressed in rad/s on three axis. Classic gyroscopes are usually composed of a spinning wheel on an axle that is free to assume any orientation. Based on the principle of angular momentum the wheel resists to changes in orientation, thereby allowing to measure values. Actually, smartphones use MEMS gyroscope sensor to detect the rotation of the device exploiting another physical phenomenon, the Coriolis force. It is a fictitious force that appears to act on an object while viewing it from a rotating reference frame.

In a mobile device, the gyroscope is a small sensor that is calibrated to give a reading of zero when the smartphone is kept on a plane horizontal surface and any change in angular rotation velocity is measured. It is a very sensitive device and it is good at detecting the spin movement. The gyroscope measures the angular velocity, which can be calculated as:

$$g(t) = \vec{g_x} + \vec{g_y} + \vec{g_z},$$

where $\vec{g_x}$, $\vec{g_y}$ and $\vec{g_z}$ are the values along the three dimensions.

Accelerometer and gyroscope are able to detect the direction of a movement, but that is a relative direction depending on the coordinates the smartphone uses. Thus, a magnetometer is needed to get an absolute direction.

Magnetometer

The magnetometer measures the strength and the direction of a magnetic field. In case a magnetometer only determines the direction of a magnetic field, it is called digital compass. Keeping the device parallel to the ground, the magnetometer provides a reference on the direction of the movement with respect to the Earth's magnetic field. The magnetic field is normally dominated by the earth magnetic field, which varies over the earth it magnitude (25-60 μT), inclination angle (0 at the equator, 90 degrees at the magnetic poles) and declination angle (-10 to +10 at most places on earth). However, when used in combination with accelerometer and gyroscope, it detects the absolute direction of the device, which is obtained regardless the position of the device. Proper calibration of the sensor is essential. Indeed, the magnetometer is sensitive to any magnetic field in the vicinity as well as hard and soft iron [21]. For indoor environments where it is not possible to obtain GPS readings, localization often relies on the magnetometer [42, 43].

Accelerometer, gyroscope and magnetometer are also called IMU (Inertial Measurement Unit).

Pressure Sensor

The number of mobile devices embedding the pressure sensor is continuously increasing. The pressure sensor can indeed be employed to measure air pressure to predict weather changes [44]. Crowd sensing methodologies for weather forecasts are highly developing all around the world. On one hand, predictions are faster than classic forecasts because smartphones act as mini weather stations located almost everywhere. On the other hand, the pressure sensor does not affect too much battery lifetime [45].

Measuring the pressure is also useful to determine the altitude. Compared to the GPS, the pressure sensor provides more accurate estimates of altitude [46] and can locate devices faster [47]. Since the GPS is not working in indoor environments, it is also a valid alternative for indoor navigation, which is an area with massive potential growth in retail and travel applications [43]. For instance, being accurate in detecting changes of altitude up to one meter, it is possible to recognize on which floor a user is inside a building [48, 49, 50]. Furthermore, in cooperation with the accelerometer and the gyroscope, the pressure sensor can recognize activities such as taking elevators, walking on stairs and changing a floor [51, 52].

Temperature Sensor

Temperature sensor indicates both the ambient and battery thermometer. Ambient temperature sensor is not one of the most popular built-in sensors in a smartphone and only few of them embed this sensor (e.g., Samsung galaxy s4 is one of the first and few smartphones to contain a thermometer). On the other hand, battery temperature sensor is always embedded in a mobile device. For this reason, it has been discovered that it is possible to measure the ambient temperature obtaining readings from the battery of the devices [53]. It exists a direct relationship exists between internal and external measures [54, 45]. Aggregating daily battery temperature readings to city level revealed a strong correlation with historic outdoor air temperature. With a mathematical transformation, the average battery temperature across a group of phones gives the outdoor air temperature. Although being uncommon, it also exist on the market standalone temperature-sensing modules that are Bluetooth-enabled and are designed to operate with smartphones. For this reason, Table 2.1 denotes the temperature sensor as an embedded sensor.

Humidity Sensor

The humidity sensor measures the air humidity, which is expressed in percentage and is computed using the absolute humidity measure relative to the maximum temperature registered for that measure. It is an embedded sensor that generates sample continuously at very low frequencies (1 Hz). In combination with other sensors, such as the temperature and pressure sensors, the humidity sensor has a potential to monitor climate conditions of relatively small environments, including working spaces like offices and laboratories and rooms [54]. Furthermore, the incoming data from millions of smartphones could potentially be utilized to create a real-world weather mapping and complex historical charts. For instance, british app developer OpenSignal has created a system that allows multiple mobile devices to provide real-time, location-specific weather reports [53].

Bluetooth

Bluetooth is a short-range wireless technology enabling low data rate communications between nearby devices. It uses short wavelengths UHF radio waves in the ISM ban from 2.4 to 2.485 GHz. Being designed for low-power consumption applications, this technology is nowadays widely adopted by the vast majority of mobile and IoT devices.

In mobile crowd sensing, Bluetooth is a well matching solution to estimate crowd density nearby, which is especially useful for collaborative sensing paradigm [55, 56]. In addition, Bluetooth is the key technology enabling connectivity between smartphones and non-embedded sensors. Note that using Bluetooth for importing data collected by a sensor that is not embedded in a smartphone permits the users to have a more accurate sensor and choose it by themselves. Moreover, it is also energy efficient in a context-aware system because the user does not have a battery consumption of the smartphone in collecting data as the case in which the sensor is built-in.

WiFi

WiFi is a wireless communication technology practically implemented in the vast majority of the mobile devices. Unlike Bluetooth, it achieves higher bit rates and higher transmission ranges, at the cost of higher energy consumption. Nonetheless, due to the limited coverage, existing WiFi infrastructure only provide intermittent connectivity for users with high mobility. Each time a user leaves the current network coverage, WiFi clients must actively discover new WiFi access points (APs) and this activity wastes precious energy because of excessive listening and scanning operations of WiFi network interface cards (NICs). Several approaches have been proposed to solve and optimise the aforementioned issue. A first solution utilizes a secondary low-power radio that communicates with peer radios on WiFi APs to find connectivity opportunities or reduce the energy consumption of data transfers. Unfortunately, this approach requires significant modifications to existing network infrastructures. A second solution predicts the availability of WiFi based on context informations. Cellular cell-tower information or together with Bluetooth contact-patterns have been used to improve WiFi prediction accuracy. However, such a context-aware approach requires extensive training based on historical information and hence it is not feasible in unknown environments.

WiFi can also be employed for localization-based service, through the WiFi positioning system (WPS). Despite GPS, it guarantees lower energy consumption and can be used for indoor localization [57]. However, WPS is not as accurate as GPS(up to 30 to 200 meters, depending on the service provider), as the minimum level of precision is nearly 20 m. Using a MCS-based approach, it becomes possible to detect presence of WiFi Access Points to build coverage maps [58].

NFC

Near Field Communication (NFC) [59] is a technology enabling smartphones and wearable devices to establish wireless communication each other when they are in close proximity(approximately 10 cm). It allows a NFC sensor chip to be recognized by simply tapping it with a NFC-enabled phone or holding the device in close proximity to it. To this date, NFC technology is more common in wearable devices than in smartphones. NFC finds application in healthcare domain [60] and for mobile payments [61]. Even if these sensors are not yet in widespread use because of the limited number of devices embedded with NFC sensors and the relative complexity and expense of producing materials with embedded chips, NFC is a growing technology platform that can be used to a variety of scenarios.

GPS

The Global Positioning System (GPS) provides location and time information. It is a space-based satellite navigation system as at least four satellites are required to compute the location in term of three dimensions, latitude, longitude and altitude. Although it does not work for indoor environments (research works show that a GPS signal is available only 4.5 percent of the time during a typical users' day), the GPS was the key enabler for location based services [62]. The GPS is very energy consuming and it is very important to switch it off when it is not necessary, indeed it can cause the battery to completely drain within a few hours. Consequently, location-based applications still cannot assume continuous and ubiquitous location access in their design because of the high energy expense of using the GPS. For such a reason, several energy-efficient techniques were investigated [63], including adaptive sampling rate methods [64]. Furthermore, often an alternative location-based service may be exploited, as WiFi or cellular tower signal, because sometimes it is better to have a less accurate position but saving a great amount of energy.

Cellular Tower Signal

Also called Global System for Mobile Communications Positioning System (GSMPS), the cell tower signal is typically employed for localizing users. Although it can not provide readings with high accuracy like GPS (the typical range spans from 70 to 200 m in urban areas), the cell tower signal is more energy efficient than GPS [65].

Ambient Light Sensor

The ambient light sensor detects the light intensity of the environment that corresponds to an approximation of the eye response to the light intensity. Being an embedded and passive sensor, the ambient light sensor is primary used for adjusting the screen brightness and the keyboard light. By all means, this was the first attempt to provide energy efficient solutions for smartphones as the screen is the major cause of battery drain. Thus, it enables easy-to-view displays that are optimized to the environment, in an effort to make the device aware of its surroundings. However, the ambient light sensor can serve many other applications. For instance, it can determine whether an user is indoor or outdoor [18] and consequently switch on or off some energy consuming sensors like the GPS. This is especially useful to design context-aware solutions that permit to switch on or off a sensor based on the position of the smartphone, e.g. in the pocket. When used in pair with the accelerometer, the ambient light sensor becomes the key enabler of modern indoor navigation systems [66, 67].

Human eye's sensitivity to light is described by the photopic curve, also called the CIE curve, which shows the sensitivity for different values of wavelength. A normalized version of this sensitivity curve is used to convert an incident optical power density (specified in µW/cm2) to sensitivity units for the human eye (specified in lux, where lux is the SI unit of illuminance and luminous emittance, measuring flux per unit area) [68]. Today most light sensors use two or more different types of photodiodes, each sensitive to a different portion of the light spectrum. By combining these photodiode outputs mathematically, each with a suitably adjusted gain, the sensor can be made to output a fairly accurate measurement of ambient brightness for the light sources commonly available. Typical values in real life are from 0.1 lux outdoor at night and 100.000 lux in sunlight. Human eye's perception of brightness is logarithmic; a lux level must increase almost ten times before the eye perceives it to be twice as bright. Thus, a similar transfer function relates the ideal percentage of display backlight brightness to relative ambient lux.

Proximity Sensor

The proximity sensor detects the presence of nearby objects and estimates the distance between the user and the mobile device without any physical contacts. The proximity sensor enables two main features. First, it helps reducing the power consumption due to the screen by turning off the backlight. Second, it disables the touch screen to avoid undesired taps, for instance when the smartphone is brought next to your ear this prevents to end a call accidentally. The proximity sensor emits an infrared light which reflects off an object back to a photodiode; thus, if there is no object, there is no reflection and no signal back on the photodiode. By the way, the infrared reflectivity of objects differs and the color has an important effect on the amount of signal reflected. It assumes that an object is close if the lights coming back are over a certain amount; if any object is present, then the touch events can be assumed to be accidental and ignored. In most mobile devices the proximity is evaluated like a boolean value, "near" or "far". A threshold is compared to a lux value, if the value is over the threshold it is near, otherwise far.

The proximity sensor is embedded in mobile devices and among various technologies, the most common implemented is the optical proximity detection [69]. It consists in transmitting infrared signal and measuring the amount of lights received back because of reflection. For this reason, Table 2.1 denotes the activity of the proximity sensor as passive. Furthermore, this sensor may be fundamental in a context-based sensing recognition(e.g., for determining if the screen of a smartphone on a table is up or down or if the device is in a pocket).

Microphone

The microphone is the most popular and well-known embedded sensor in smartphones as it is fundamental for calls, the primary functions of telephony. When used for sensing, the microphone is a precious source of information. For example, the majority of places and sites have specific sound patterns. As a result, this information becomes fundamental to understand users' context in everyday life [70, 71]. Moreover, also activities such as conversation, music, traffic noise and ambient sound have a unique fingerprint that can be categorized [72, 73, 74]. Having this information, it becomes possible to instantiate users' profiles that change according to the context. To illustrate, adjust the volume of the smartphone according to the environment noise. Furthermore, when microphone is used as a sensor for environmental noise, it is fundamental to decide an efficient sampling period. For example, the period can be set according to the remaining battery charge of the users and the needs of the collector about quality of signal and sensed data. This sensor is strictly related to the concept of privacy, indeed it can be exploited only under the explicit permission of the user, who has a primary and conscious role in the sensing process.

Camera

Similarly to the microphone, the camera is certainly one of the most popular embedded sensors in smartphones. Its typical use consists in capturing images and recording videos. In phone sensing-based applications, the camera is very popular and finds application in a broad range of domains such as healthcare [75, 20], intelligent transportation systems [76, 7, 77] and environmental monitoring [24, 24]. However, it should be noted that image processing is typically computational heavy if performed locally on smartphones. Other promising use cases are indoor navigation [78, 79] and augmented reality-based translation systems [80].

In contrast to other smartphone sensors, but similarly to the microphone, it can be exploited only under the explicit permission of the user. Moreover, it is needed an active action of the users, who must take a picture with their smartphone of the target object and upload it to a central service where all the collecting data will be analyzed. For instance, the image could result useless if not focused or not taking the interested subject. Furthermore, the view of the camera is often obstructed, as when a mobile device is brought into a pocket.

Air Monitoring Sensor

Air pollution is certainly one of the most important worldwide concerns nowadays. MCS systems relying on large user participation have a great potential for air monitoring. However, most of the air monitoring sensors available are non-embedded, but standalone devices. As a result, special incentives should be put in place to motivate users buying and using such devices. Typically, air monitoring sensors are equipped with Bluetooth [81], but it exists implementations that require a wired connection such as the GasMobile hardware [12]. It is a sensor for consumer and industrial applications to monitor gases like carbon monoxide, oxygen, ammonia, fluorine, chlorine dioxide and others. The dust sensor [13] is a particular and very promising prototype for air monitoring, with the possibility to consult or share the data to the community. This sensor, differently from the others concerning the food allergens, could be deployed also in workplaces to guarantee the workers' safeness without spending a great amount of money, but having a simple sensor easily accessible by workers.

Gluten Sensor

The gluten sensor is a non-embedded sensor which makes possible for people with celiac disease to monitor and share food properties [11]. The device only works in pair with a smartphone to track statistics. It is an affordable handheld device that allows users to quickly and easily check foods for gluten. It consists of two parts: a sensor pod and a number of single-use disposable testing units that users dip into their food. Once a sample is collected, the testing stick is put into the pod, the device analyzes it and then send the test result to the smartphone via Bluetooth. There is also the possibility to share results amongst users, to be informed about which restaurants or dishes are safe for those with food allergies or particular food sensitivity. This is one of the best example about health care and food sensing, collecting and sending data and also sharing in cloud. Thus, other devices may result useful for a wide range of different allergens in the same context, such as peanuts. This is due to everyday people are becoming more aware and more concerned about what hidden things are in the food they are eating. Furthermore, this is the typical example of an application that does not need any incentives, as the people are interested in joining the community.

Radiation Sensor

Similarly to the gluten sensor, the radiation sensor is non-embedded. It detects nuclear radiation samples and complex operations such as A/D conversion, filtering and threshold comparison are offloaded and executed by a smartphone. The idea originated by the users need in having cheap and affordable mobile radiation detectors, which appeared to become urgent after the nuclear disaster of Fukushima [14].

Table 2.1: Classification of Mobile Devices Sensors. P denotes passive sensors, A denotes active sensors.

Sensor	Acronym	Description	Purpose	Embedded	Activity
Accelerometer	ACC	Measures the acceleration of the device	Orientation and inertial	✓	P
Gyroscope	GYR	Detects the orientation of the device and measures the angular velocity	Orientation and inertial	✓	P
Magnetometer	MAG	Measures strength and direction of a magnetic field	Orientation and inertial	✓	P
Pressure Sensor	PRE	Measures the air pressure	Weather	✓	A
Temperature	TEM	Measures the temperature of the environment	Weather	✓	A
Humidity Sensor	HUM	Measures the humidity of the environment	Weather	✓	A
Bluetooth	BLT	Allows to find and connect to other devices	Communication	✓	A
WiFi	WIFI	Scans available networks and detects the approximate position of the device	Communication/Localization	✓	A
NFC	NFC	Enables communication with devices in close proximity	Communication	✓	A
GPS	GPS	Provides information on time and position	Localization	✓	A
Cell Tower Signal	CTS	Provides a very approximate position of the device	Localization	✓	P
Ambient Light	AML	Measures the light intensity of the environment	General Purpose	✓	P
Proximity Sensor	PRO	Detects the presence of nearby objects	General Purpose	✓	P
Microphone	MIC	Detects sounds and noise of the environment	General Purpose	✓	A
Camera	CAM	Captures pictures and videos	General Purpose	✓	A
Air Monitoring	AIR	Monitors air quality and gas concentration	Specialized	✗	A
Gluten	GLU	Checks foods for gluten	Specialized	✗	A
Radiation Sensor	RAD	Measures nuclear radiation	Specialized	✗	A

Chapter 3

Mobile Crowd Sensing

Mobile Crowd Sensing (MCS) has gained significant attention and is becoming a new appealing paradigm for sensing. For data collection, MCS systems rely on contribution from mobile devices of large number participants, or a crowd. Smartphones and wearable devices are deployed widely and already equipped with a rich set of built-in sensors, making them an excellent source of information. The ubiquitous diffusion of mobile devices along with the rich set of built-in sensors they are equipped with are certainly the two main key enablers leading to the success of MCS paradigm. Accelerometer, GPS, camera and microphone are only a representative set of sensors equipped in mobile devices. Although MCS is an emerging paradigm, a number of MCS applications relying on mobile devices sensors have already been developed in different scenarios, including healthcare, environmental monitoring, public safety and intelligent transportation systems such as traffic monitoring and management [82, 83]. All these applications suit very well urban scenarios. As a consequence, MCS is an essential solution for building smart cities of the future, which aim at using ICT solutions to improve management of everyday life of their citizens [5].

Mobility and intelligence of human participants guarantee higher coverage and better context awareness, if compared to traditional sensor networks. On the other hand, individuals may be reluctant to share data for privacy concerns. For this, the research community in MCS has put lot of effort in developing incentive mechanisms to foster user participation [84, 85] and in investigating privacy issues [86, 87]. However, privacy is not the only barrier limiting user participation because sensing is costly. The reason is twofold: from one hand mobile devices are battery constrained, hence it becomes important to exploit such resources appropriately, i.e. not perform unnecessary sensing operations. On the other hand, collected information often requires to be delivered to a central collector. Communication technologies such as 3G/4G, WiFi or Bluetooth affect battery lifetime differently [88, 89] and have different monetary costs for reporting.

Despite of the growing interest in the research community, MCS solutions remain largely uncategorized. This chapter develops a detailed taxonomy to classify MCS applications, methodologies and architectures and the most common sensing elements available in today mobile devices. The objective is not only to classify, analyze and consolidate past research, but also to outline potential future research direction.

3.1 Background on Mobile Crowd Sensing Systems

MCS systems rely on sensors and communication interfaces embedded in commonly used mobile devices such as smartphones, tablets and wearables [90, 91]. Although being battery constrained, mobile devices are nowadays powerful, having computing, communication and storage capabilities. Indeed, they are essential for our daily activities, including business, communication, social activities and entertainment [92, 93]. According to Gartner statistics, the number of worldwide smartphones sales in 2015 was 1.4 billion units [94]. Wearable devices are increasing in popularity as well. The number of wearables shipped in 2014 was 70 millions, which is projected to reach 91.3 millions in 2016 [95]. Smart watches, glasses, rings, gloves and helmets are the most popular wearable devices currently available on the market, which is projected to rise up to $ 30.2 billion by 2018 [96].

The term *mobile crowd sensing* was first introduced by Ganti et al. [90] and indicates a more general paradigm than mobile phone sensing. Guo et al. in [91] give a definition that clearly highlights this difference: *"MCS is a new sensing paradigm that empowers ordinary citizens to contribute data sensed or generated from their mobile devices, aggregates and fuses the data in the cloud for crowd intelligence extraction and people-centric service delivery".* To operate efficiently, MCS systems require participation and contribution of a large number of users. Although entire communities can potentially benefit from such a contribution, singular person may be reluctant to participate, being selfish or having privacy concerns. To easy this burden, in the last years the research community has put lot of effort in developing proper incentive mechanisms [97] and in investigating privacy issues [84, 85, 87, 86].

The capillary spread of smartphones and wearables along with the rich set of built-in sensors these devices are equipped with are certainly the main key enablers leading to the success of MCS paradigm. Accelerometer, gyroscope, GPS, microphone and camera are only a representative set of sensors that facilitated the development of a number of applications in a wide range of scenarios, including health care, environmental and traffic monitoring and management. Many applications using smartphone sensors have been already developed and are currently in use, as surveys on phone sensing systems classify them [82] and [83].

To illustrate few representative examples, HealthAware [75] and DietSense [20] foster healthy eating by collecting images of consumed food and inspect daily user-activity by extracting context information such as time and location where food was consumed. For this purpose both applications use accelerometer, GPS and microphone. Nericell [76] and the Pothole Patrol [7] monitor road surface and traffic conditions. NoiseMap [24] and GasMobile [24] monitor noise and air quality respectively.

MCS can greatly improve citizens everyday life and provide new perspectives to urban societies. MCS is an essential solution for building smart cities of the future, which aim at using ICT solutions to improve management of everyday life of their citizens [5]. The Internet of Things (IoT) paradigm is the candidate solution to provide a simple infrastructure to foster the smart cities. For this, it is required to deploy sensors equipped with communication capabilities widespread [6]. According to Gartner statistics, during 2015 the number of connected objects in smart cities exceeds 1.1 billion and it is expected to grow up to nearly 10 billion by 2020 [98]. In such a context, active participation of citizens can improve spatial coverage of already deployed sensing systems with no need of further investments. MCS leverages human intelligence, which has a deeper context understanding than traditional sensor networks. For example, having human involved in detection of free parking spot detection provides more accurate performance [90, 9]. To illustrate, ParkSense detects vacant parking spots using WiFi scans of smartphones [8]. Parking is only one of the possible city services where MCS can play a fundamental role thanks to its unique features. In addition, other potential applications are smart traffic management [99, 100] and environmental monitoring, including air [81, 12] and noise quality [101, 24].

This chapter provides a taxonomy, definitions and comprehensive analysis of the available mobile crowd sensing solutions. During the past five years, the research community proposed a number of sensing architectures and focused on the analysis of incentive mechanisms and privacy issues as well as the reliability and trust of data collection process. With the sole exception of [102], this vast amount of work remains uncategorized, with many of the core paradigms undefined. To illustrate, for example, there is no consensus on the term "opportunistic sensing". According to Ganti et al. [90] "opportunistic sensing" is defined as "*On the other hand, opportunistic sensing is where the sensing is more autonomous and user involvement is minimal (e.g. continuous location sampling)*". However, Khan et al. [83] state that opportunistic sensing requires no user involvement at all, since the decisions to perform sensing is a prerogative of the device itself. Finally, Han et al. [103] enlarge previous vision of opportunistic sensing in the context of single user involvement and they describe opportunistic sensing as a paradigm enabling cooperation among smartphones. Typically, both terms "opportunistic" and "participatory" sensing remain consider under the common umbrella of MCS [90, 83, 104, 103]. In other cases, both "mobile crowd sensing" and "participatory sensing" are used interchangeably [99]. Other times, both "mobile crowd sensing", "participatory sensing" and "opportunistic sensing" are synonyms [105]. With this plethora of definitions, the objective of this chapter is to simplify understanding of the current definitions and available techniques and solutions in the field of MCS.

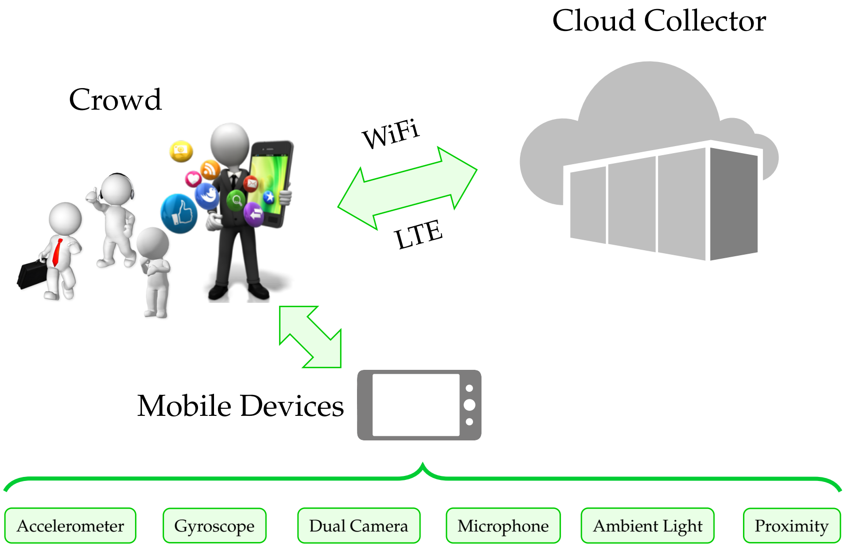


Figure 3.1: Cloud-based MCS system

3.2 Taxonomy

To easy and simplify the understanding of MCS this section elaborates a novel taxonomy that provides a clear organization of the state-of-the-art. Figure 3.2 illustrates the proposed taxonomy, which is a layered taxonomy. The proposal provides a hierarchical view, where on the first level the classification groups the works according to their *application target* (Section 3.2.1). On a second level, this category captures in detail properties of MCS systems and analyzes works in relation to their application target. Specifically, this section considers the methodologies through which the sensing activity is performed, for example how users join MCS systems (*sensing activity* category), how tasks are assigned and executed (*task* category) and the properties of reported data (*data property* category) In addition, it also analyzes the sensing elements used per application (category *type of sensor*). Despite the taxonomy and the classification illustrated in Section 2, in this section the focus is not on the properties of the sensors, but a more detailed analysis on motivations and goals of using sensing elements for each application target is provided. Next, Section 3.3 overviews the available literature exploiting this layered taxonomy.

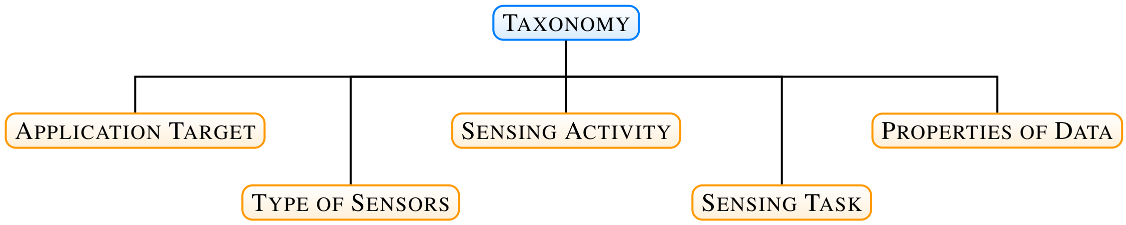


Figure 3.2: Mobile crowd sensing taxonomy

3.2.1 Application Target

The application target is the key layer of the taxonomy. Any element of such category corresponds to a different application, for example environmental monitoring. Having this key layer allows to unveil particular properties of MCS systems. For example, health care applications usually employ general purpose sensors while environmental monitoring applications rely on specialized sensors very often. Table 3.1 briefly summarizes all the applications that are described in more details in the following.

Table 3.1: Classification of Application Targets

Target	Acronym	Description
Environmental Monitoring	EM	Measure the quality of the environment
Health Care	HC	Measure biological parameters influencing human health
Intelligent Transportation Systems	ITS	Improve traffic forecasting and public transportation design and management
E-Commerce	EC	Analysis of goods price for offline markets
Public Safety	PS	Measures and initiatives for public safety
Human Behavior	HB	Analysis of human activities and emotions while using mobile devices
Mobile Social Networks	MSN	Exploit social interactions among users such as friendship
Other	OTH	Other scientific contributes not included in previous categories

Environmental Monitoring (EM)

Monitoring the environment is very important for human health. This is especially true in cities, where the majority of the worldwide population lives. To this end, MCS is the ideal solution to improve strategies and policies that aim at understanding the current status of the environment and provide forecast analysis. Ubiquity and mobility of mobile devices allow fine-grained analysis of environmental phenomena like air pollution and noise. For example, PEIR [106] analyzes on a per-user basis how transportation choices simultaneously impact on the environment and which is the risk-exposure for the individual.

Air Quality monitoring Clean air is fundamental for human health. However, air pollution in cities is a major issue today [81]. Indeed, atmospheric pollutants are responsible for a variety of respiratory diseases and can be cause of cancer if individuals are exposed for long time periods [12].

Noise Monitoring Noise pollution in urban areas is a well-known problem affecting every-day life of citizens. For the authorities it is very important to monitor this phenomena to adopt proper strategies and countermeasures [101]. Creating noise maps [24] is a common methodology to analyze noise pollution. However, conventional methods are often very expensive, inaccurate and rarely up to date. Leveraging citizen participation certainly lowers this burden to provide open and inexpensive noise maps and data graphs [107] .

Health Care (HC)

Health care consists in diagnosing, treating, and preventing illness, diseases and injuries. This category classifies all MCS systems which make use of data for individuals and community heath care purposes, including food and fitness monitoring, elder support and assistance.

Intelligent Transportation Systems (ITS)

Intelligent transportation systems (ITS) rely on IT solutions with the goal of providing innovative services, improving cost effectiveness and efficiency of transportation and traffic management systems. For example, wind warning systems alert drivers approaching bridges in case of dangerous wind conditions. The most well-known application of ITS systems is in monitoring traffic and road conditions. Detecting traffic conditions through MCS paradigm provides to the citizens estimations of the time needed to reach some place [76] or about next bus arrival [108]. As for road monitoring, the combined use of accelerometer and GPS helps detecting holes and bumps [7].

E-Commerce (EC)

Websites enable users to track and compare price information, however this is very difficult for offline markets. MCS can help tracking price dispersion, which is the difference in terms of price of the same good among different vendors [109].

Camera and GPS used in combination can track fuel prices, which are later compared with information gathered by other users to detect most convenient petrol stations [19].

Public Safety (PS)

Nowadays the public safety is one of the most important and challenging issues for general public and administrations, which includes protection and prevention from consistent damages, injuries or generic dangers. Typical examples to take into account for public safety are burglary, trespassing, harassment, inappropriate social behaviour, flooding or earthquakes. For instance, it is possible to evaluate the safety of citizens exploiting data collected from geosocial networks and relating them with crime [110]. In other words, public safety is responsible for taking appropriate countermeasures about man-made crimes, incidents like crash accidents or natural disaster like flood.

Human Behaviour (HB)

Mobile devices have a potential to classify human behaviour. For example, for psychologists it is important monitor patients without asking to recall events after long time. The use of mobile devices easy this burden [111]. Although sensors can efficiently perform activity recognition, some human activities are complex and require machine learning techniques to infer particular moments of people's life [112].

Mobile Social Networks (MSN)

Mobile Social Networks are communication systems that rely not only on human behaviours and activity recognition but also on the social needs of the users [113], exploiting mobile devices. In this category the most fundamental aspect is the collaboration between users, who share with the other participants the data sensed through smartphone sensors, such as recognized activities and locations.

Others (OTH)

This field classifies all the scientific contributions that does not deal with the issue of specific user behaviour and does not have a particular target application. These are useful and contribute in MCS area of research in other general fields, such as energy efficiency, incentives, utility maximization, workload allocation, data collection, frameworks, platforms, etc.

3.2.2 Type of Sensors

This category investigates how the most common sensing elements analyzed in Section 2 are used per application target. They are summarized in Table 2.1.

3.2.3 Sensing Activity

The *sensing activity* category analyzes the methodologies and procedures used to perform sensing. This category includes strategies that are used to involve users in the process, who is responsible for taking sensing decisions, the policies adopted for data reporting and whether collected data can be useful for single individuals or to a community. Figure 3.3 illustrates the terminology considered for this category in more details.

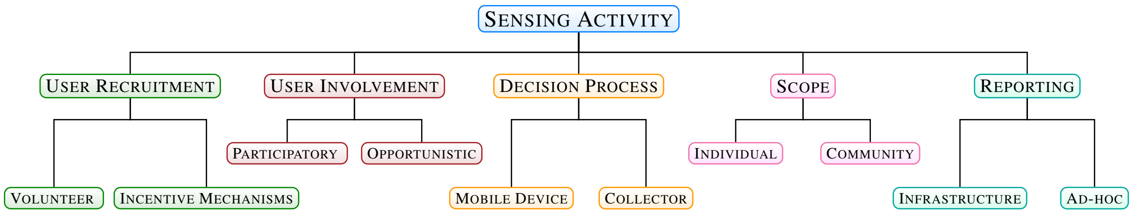


Figure 3.3: Sensing activity taxonomy

User Recruitment

The category *user recruitment* denotes the process in which users join the MCS system and it can be based on volunteer participation or through incentive mechanisms. This classification considers strategies and mechanisms used to create a MCS system, which is to motivate users to participate. The strategies are strictly related to the application target of the MCS system. For instance, for a healthcare application that collects information on food allergies (or gluten-free [11]), people affected by the problem typically volunteer to participate, being very interested. However, if the application target is more general and does not match well users' interest (e.g., noise monitoring [24]), the best solution is to encourage user participation through incentive mechanisms. It should be noted that these strategies are not mutual exclusive. That is, it could be perfectly possible having an application for which users volunteer and receive a reward for their contribution. Such solution could be suitable in case users should perform additional tasks (e.g., sending more data) or in particular situations (e.g., few users contribute from remote area).

Volunteer In volunteer-built MCS systems users are willing to participate and contribute on spontaneous basis and they typically do not expect to receive a reward. A framework for recruitment is proposed in [133], where the organizers are allowed to choose well-suited participants according to some features including their habits and data collection is based on geographical and temporal availability. Apisense is another example of platform helping organizers of crowd sensing systems to collect data from volunteers [134].

Incentive Mechanisms Incentives are required when users do not have a strong motivation in participating in the MCS system. It exists many strategies to stimulate participation [97, 87, 135]. Zhang et al. [84] propose to classify the existing works on incentives into three categories: entertainment, service and money. In the first category they consider methods that stimulate people by turning some sensing tasks into games, so users can enjoy while participating in sensing. The second category is about service-based mechanisms, which consist in remunerating personal contribution with system service. Finally, monetary incentives methods provide money as reward for users' contribution.

User Involvement

Understanding how users are involved in the sensing process is very important to adopt proper incentive mechanisms. Indeed, different application target may require users to be actively engaged in performing specific tasks. In other words, after having received a task, the user needs to perform some activity, for example turning on the microphone to report audio samples. However, there are cases where it is more beneficial having users collecting data opportunistically. This occurs when no specific task is assigned in origin or the users do not have to perform specific operations consciously, being sensing operations triggered automatically by the mobile device. *Partecipatory* and *opportunistic* denote respectively the cases in which users are actively engaged or not in the sensing process respectively.

Sometimes it is difficult to define exactly if a sensing paradigm is participatory or opportunistic and also in literature these definitions are often misleading. For instance, the authors of a system that predicts bus arrival time [108] define the user involvement as participatory, considering that users decide actively to participate in the sensing process. Actually, in the proposed taxonomy it is an opportunistic involvement because the user is unconscious in respect to the sensing process, which exploits accelerometer and microphone to detect when a user is on a bus ("am I on a bus?") and to collect data. Just to clarify, in the proposed classification the sensing would have been participatory if the user had communicated actively to be in a bus and to share the position.

Participatory MCS systems relying on participatory policies require users to be actively involved in sensing operations as they are asked to perform some tasks. To this end, users are responsible to decide when, what, where and how to perform sensing (e.g., to take some pictures with the camera due to price comparison [115], or record audio due to noise analysis in cities [101, 136]). Under such a paradigm, the users are free to accept or not incoming tasks and they are rewarding on the basis of the number of accepted tasks and the quality of accomplishment. An advantage of this approach is that complex operations can be supported by exploiting the intelligence of the users. Indeed, they can solve the problem of context awareness and consciously meet the application requests in a very efficient way. For instance, a person who wants to get involved in collecting a sample of noise pollution in the neighborhood simply takes the smartphone in the hand to accomplish the task. On the other hand, a problem strictly related to this advantage is that the quality of data is totally dependent on participant enthusiasm and ability to reliably collect sensing data. Applications are best suited to the participatory model when they have a collection of interested users whose size is at least as large as number of sensors required to map a phenomena efficiently (e.g., celiacs [11]).

Opportunistic MCS systems relying on opportunistic sensing paradigm do not task users specifically, minimizing their involvement in the process. This means that users have only to declare their interest in contributing data. Then, all the decisions of what, where, when and how to perform sensing are demanded for example to an application running on the mobile device. To this end, context-aware sensing becomes very important for opportunistic sensing paradigm. Indeed, capturing images when the smartphone is in the pocket not only does not bring any utility for the system, but also has a cost in terms of energy for the device. As a consequence, context recognition has been largely investigated in the recent years. For example, to detect road conditions, accelerometer and GPS readings are obtained only when users move by car [7]. Other fundamental challenges that need investigation are providing sensing coverage when sensor mobility is uncontrolled, ensuring consistent sensor calibration and protecting custodian privacy.

- **Active.** Tasks are triggered by the smartphone when some context or a certain situation in the user surroundings that is requested by the task authority is recognized. This is the approach this thesis will take into account in chapter 4 presenting a framework for data collection.

- **Passive.** Tasks are accomplished by the software automatically. Usually the data collector decides a sampling rate for sensing the data.

Decision Process

The category *decision process* analyzes the sensing activity on the basis of who is responsible for taking decision to collect data. To illustrate with an example, when the mobile devices take decisions, samples can be collected upon meeting specific condition such as in advantageous context. This category identifies two main paradigms (see Fig. 3.3). From one hand *mobile device* or user can take decisions locally. This is a distributed paradigm and may require coordination among the devices if users are tasked. On the other hand, the crowd sensing *collector* can be designed to be responsible for taking sensing decisions. The collector decision process has a centralized view of the amount of information already collected and therefore can task users or demand for data in a more efficiently manner.

Mobile Device A *mobile device* or a user takes sampling decisions locally. In the latter case, the single individual decides when, where, how and what to sample, e.g. taking a picture for sharing the costs of goods [114]. It indicates the participatory paradigm. When devices take sampling decisions, it is often necessary to detect the context in which smartphones and wearable devices are, as the opportunistic paradigm. The objective of this local approach is to maximize the utility of data collection and minimizing the cost of performing unnecessary operations.

Collector When the *collector* takes sensing decisions, it has to communicate the decisions to the smartphones. Centralized decisions can fit both participatory and opportunistic paradigms. If the requests are very specific, they can be seen as tasks by all means and indeed the centralized decision paradigm suits very well the participatory sensing approach. However, if the requests are not very specific, but contain generic information such as “send more audio samples”, the opportunistic paradigm is exploited.

Scope

The *scope* category identifies who can benefit from data collection and processing. This category includes single individuals and communities. In the first case, the main interest is personal. An example of this kind is sensing data to track biological parameters for fitness and wellness. However, it exist many cases where communities can take advantage from sensing. For example, monitor air quality or road conditions. When the scope is community-based, large user participation is needed in order to have many samples to better characterize the phenomena.

Individual The scope of sensing activity is considered to be individual when it concerns a process focusing on personal interests. Data is sensed from the crowd for individual needs and each user can receives a sort of feedback that may result helpful for some activities or habits (e.g., monitoring daily physical exercises and diet [75]). For instance, in healthcare domain, the users sample data to have a feedback about their health condition [117] or they can better take care about their food allergies (e.g., celiac people [11]).

Community When the benefits of sensing activities concern entire communities, large user participation is required to better characterize the phenomenon. In this context the utility given by the contribution of a single user is marginal. For example, while monitoring air pollution [81], a single individual does not have an immediate advantage while she contributes samples.

Reporting

Once data is collected from the mobile device, it is typically delivered to a central collector responsible for processing. This category considers two possible methods to report data, namely *infrastructure* and *ad-hoc*-based. Infrastructure-based reporting operates through cellular or WiFi networks. Ad-hoc-based reporting involves opportunistic communications between the devices, using technologies like Bluetooth and WiFi-Direct. With this paradigm samples need to hop many times before reaching the collector. For such a reason, ad-hoc reporting should not be used for application having strict constraints on latency.

Infrastructure Reporting data through infrastructure-based communications is a more energy-demanding task. Indeed, technologies used for opportunistic communications like Bluetooth and WiFi-direct are more energy-efficient than cellular (3G/4G) and WiFi networks. However, having infrastructure-based communications enables fast reporting of urgent data. For this, cellular connectivity should be preferred as WiFi connection is not always available (e.g., building a real-time map for monitoring availability of parking slot [9]).

Ad-hoc This category includes data forwarding utilizing intermittent connections with short-range radio communications, such as Bluetooth and WiFi-Direct. If Bluetooth is used, it is possible to exploit a collaborative approach with other devices in the neighborhood. This can also enable coordination for distributing task execution.

3.2.4 Sensing Task

In MCS systems, the notion of *sensing task* is extremely important. Both participatory and opportunistic sensing paradigms (see Section 3.2.3) make use of tasks to achieve the target. In the participatory paradigm they are explicitly revealed to users, whereas in opportunistic sensing they are implicit. This means that the system itself adopts mechanisms that encourage users to contribute to achieve a given target, leaving to the single individual the freedom to decide the level of contribution. Fig. 3.4 illustrates the proposed taxonomy to characterize sensing task properties.

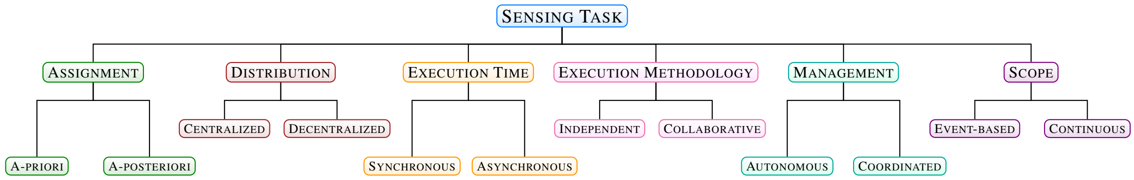


Figure 3.4: Sensing task taxonomy

Assignment

Task *assignment* describes the process of attributing jobs to users. On the basis of the time occurrence of the phenomena to capture, two main strategies can assign tasks, a-priori and a-posteriori. Note that task assignment should not be confused with task distribution, which is considered as a different category.

A-priori Tasks should be assigned *a-priori* when the objective to be achieved is already clear before starting the sensing campaign. To illustrate with an example, monitoring a phenomenon like temperature or air pollution [81] fall in this category.

A-posteriori Tasks can be defined to assigned *a-posteriori* when the main interest is not monitoring a phenomenon, but receiving a feedback on an already occurred event. This approach is certainly helpful in public safety context, for example to receive information from users that assisted to a crash accident. In addition, in many applications exploiting social networks tasks are decided only after users have already performed sensing [122, 123].

Distribution

Tasks can also be classified according to the way in which they are distributed among the users. Following the model of Pournajaf et al. [87], *distribution* introduces three main categories: centralized, decentralized and hybrid. These categories highlight the entity responsible for task assignment. Note that this classification differs from *task assignment* classification, which focuses on the instance of time tasks are generated.

Centralized The distribution is *centralized* when the tasks are created by a central authority or a tasking entity and they are directly assigned to all the participants. Users do not re-distribute tasks among them, but they only have to perform the assigned job. Typical centralized task distribution involves environmental monitoring applications such as detection of ionosphere pollution [138] or nuclear radiation [14].

Decentralized When the task distribution is *decentralized*, each participant becomes an authority and can either take the decision to perform the task or to forward it to other users. This approach is very useful when users are very interested in some events or activities and the purpose is not to monitor a phenomenon. A typical example is Mobile Social Network or Intelligent Transportation Systems, to be aware of public transport delays [108] or availability of parking spots [9].

Hybrid The *hybrid* distribution has features of both previous categories. For example, tasks can be assigned by a central authority for new participants joining the system. Later, users act as peers receiving tasks from other peers in a collaborative manner.

Execution Time

Execution time of tasks can be synchronous or asynchronous. This category considers the cases in which the participants have to start the sensing process all together in the same moment or not.

Synchronous This category analyzes the case in which the users start together the sensing activity. For synchronization purposes, the participants can communicate each other or receive an exact time from a central authority. For instance, LiveCompare [115] compares the price of goods and the users should start sensing synchronously, otherwise the comparison does not provide meaningful results. This approach is strictly related to the a-posteriori task assignment.

Asynchronous The execution time is *asynchronous* when the users perform activity independently. This approach is used for monitoring, because the ultimate goal is to receive a certain amount of data regardless the start time of the sensing campaign. Examples of asynchronous task assignment are noise [24] and air pollution [12] monitoring.

Execution Methodology

Execution methodology analyzes whether tasks execution is performed in collaborative manner by the participants.

Independent Task execution methodology is *independent* when each mobile device accomplishes the requested task independently from other devices. If the device finds itself in a bad context or is running out of battery, the tasks can be rejected.

Collaborative Methodology is *collaborative* when the smartphones execute the requested tasks with the help of other mobile devices. Despite independent execution, if the device is in unfavorable conditions, the task can be forwarded to other users.

Management

Task *management* describes how the system can be categorized considering the task allocation scheme [139]. In other words, it determines if participants must be able to provide the data requested from the collector by themselves or they can manage the task to other users involved in the sensing process(e.g., a mobile device cannot sense the temperature of a place and it manages the task by itself asking to another user in the surroundings).

Autonomous A management is autonomous when a user does the task by his own, in a way totally independent from the other ones. If a user does not succeed in accomplishing the task, it is not done by another one. It will be the central authority, or the collector, to give the task to other participants.

Coordinated Task management is coordinated when users cannot accomplish a task or find someone who can do it in a better way and make direct queries to other users. A typical example of coordinate allocation is Micro-Blog [124], in which every time a particular content is not available on the map, the participants can mark out a geographic position and make direct queries to all smartphones that are located close to that location.

Frequency

This category analyzes how often a task has to be executed. Some types of tasks can be triggered by event occurrence or because the device is in a particular context. On the other hand, it exists tasks that have to be performed continuously.

Event-based The frequency of task execution is *event-based* when data is sensed after an event has happened. Events can be the occurrence of a particular situation or a given context activity (e.g., users moving outdoor, or getting on a bus). As a result, event-based task execution is not regular, but it requires a decision to trigger the process. In participatory sensing paradigm, users are typically aware of their context and can perform directly the task. For example, monitoring a community interest (e.g., food allergies [11]) or sensing data after an event, such as taking pictures after a car accident or a disaster (e.g., earthquake, flooding). In opportunistic sensing, tasks can be event-based upon recognition of the context (e.g., user getting on a bus[108]).

Continuous This category classifies the tasks that are accomplished regularly and independently by the context of the smartphone or the user activities. The data collection continues until there is a stop from the central collector (e.g., quantity of data is enough) or from the user (e.g., when battery level is low). In a continuous sensing it is very important to set a sampling period that should be neither too low nor too high, to result a good choice for data accuracy and in the meanwhile not too energy consuming. For instance, air pollution monitoring [81] must be based on a continuous sensing to have relevant results and should be independent from some particular event.

3.2.5 Properties of Data

Sensors generate different types of data and analyzing data properties is very important. For example, images captured after a crash accident have a different priority than temperature samples. The urgency of data delivery is very strict in the first case while it is delay-tolerant in the second case. Moreover, the amount of data to be transmitted is very different: images are heavy in size if compared to temperature readings. Typically, applications make use of different types of data simultaneously. For example, Dong et al. [19] developed an application to detect fuel prices in gas stations using both pictures and location records. While approaching gas stations, an algorithm extracts the price from pictures recorded through the camera. The location of the gas station is associated to each price thanks to the GPS. Fig. 3.5 illustrates the proposed taxonomy for this classification. In addition to distinguish different data *types*, the taxonomy investigates *delivery* strategies and *granularity* of data. It should be noted that the categories *delivery* and *reporting* (see Section 3.2.3) differ because they investigate respectively the delivery urgency based on the type of data and technologies used for reporting.

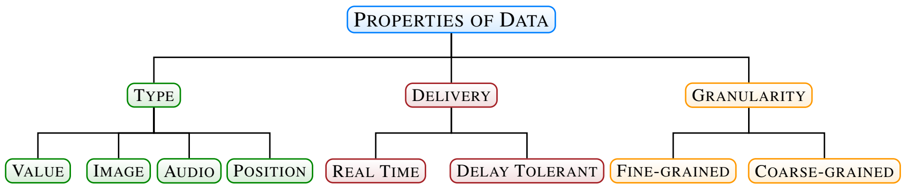


Figure 3.5: Properties of data taxonomy

Type

Different sensors generate a different *type* of data. To be as general as possible, this category defines types to be representative. Indeed, the same sensor can generate data in form of different types. For example, in NoiseMap [24] the microphone does not record an audio file as usual, but it measures a value in dB corresponding to the level of noise in an given location.

Value The term *value* indicates data types for which it is associated a numerical value (e.g., a noise level or raw data of accelerometer) or a classification label (e.g., user moving or not, car or bus movement detection, noise of something particular [140]).

Image This subcategory considers both images and video files. In the first case, data is obtained using a built-in camera in the smartphone, whereas video are usually recorded exploiting both camera and microphone.

Audio *Audio* files are sampled with built-in microphone. It is important to distinguish a precise audio signal recorded and classified as audio with respect to other applications where the microphone detects and classify level of noise or voice.

Position The *position* corresponds to the location of the device at time of sensing. To each position record, information on latitude and longitude is always given whereas the altitude is typically optional.

Delivery

The *delivery* category analyzes the urgency data should be sent to the collector. There are two main strategies: real-time and delay-tolerant.

Real Time *Real time* data delivery occurs when the collector needs to receive and process data very fast. Typical applications with such requirements are those related to emergency. In this case, cellular and WiFi technologies should be preferred for communications. Consequently, an important aspect that needs more investigation is the problem of congestion control in large wireless networks where nodes periodically broadcast time-critical information [141]. A proposed solution could be to implement scheduling and queue management techniques, in which a queue is maintained with only the latest status packet of each source, overwriting any previously queued update from that source [142].

Delay Tolerant When the purpose of the application target does not have strict requirements in receiving and processing data. Under such scenario, opportunistic communication technologies can be a good option.

Granularity

The *granularity* helps analyzing the level of detail data is sensed. This category identifies two main levels: *fine-grained* for high precision measures and *coarse-grained* (see Fig. 3.5). For instance, the notion of position is fine-grained when it provides exact information on latitude,longitude and altitude. However, it can also be coarse-grained when the value provided is just an approximation, for example when it is used to distinguish outdoor from indoor positions. Applications can use fine and coarse levels of details simultaneously.

Fine-grained Data granularity is *fine-grained* when it is very detailed and precise. To illustrate some examples, readings obtained from the accelerometer and GPS and pictures can be attributed to this category.

Coarse-grained Data is *coarse-grained* if obtained with less levels of detail or in low resolution. For example, the GPS can also be used only to detect indoor or outdoor position.

3.2.6 Privacy

MCS involves the collection of detailed information from users' mobile devices during task management processes [87]. It implies the possibility that collected data could compromise not only participants' privacy but also unconscious people in the surroundings (e.g., recording an audio in which somebody is speaking). Several privacy concerns could be taken into account. For instance, the identification or disclosure of sensitive attributes increases vulnerability and subsequently reduces participation [86]. Kantarci et al. consider trustworthiness of crowdsourced data an important challenge in S^2aaS (Smartphone Sensing as a Service), as maliciously altered data can be misleading for the S^2aaS customer [143]. Nonetheless, considering these aspects is not the scope of this thesis' analysis. This section aims only at classifying works considering if they have privacy constraints or not. For instance, when the task is monitoring air quality there are no privacy constraint. On the contrary, when it is necessary to use sensors as microphone or camera, the data collection process is strictly related to the privacy, overall if the user is in a public space (e.g., taking and sharing pictures about a car accident for public safety in which many people are involved).

3.3 Overview on the Architectures

This overview presents some of the most cited works about MCS systems and classifies them exploiting the taxonomy proposed in 3.2. First, this section shortly explains and distinguishes them according to the taxonomy of target application proposed in 3.2.1. Then, specific tables classifies MCS architectures according to type of sensors, sensing activity, sensing task and properties of reported data.

Environmental monitoring

Mahali Project [116] is one of the newest work in environmental monitoring. Recent discoveries of signature in the ionosphere related to earthquakes and tsunamis suggest that ionosphere may be used as a sensor that reveals Earth and space phenomena [138]. Consequently, Mahali exploits GPS signals to enable a tomographic analysis of the ionosphere, which is seen as a global earth system sensor. Furthermore, it is built to support well different configurations of MCS, for instance delivery can be done with infrastructure or opportunist approach, data collection in synchronous or asynchronous mode and a potentially good educational incentive mechanism is proposed to recruit and involve people in the scientific contribution. The Personal Environment Impact Report (PEIR) [106] exploits location data sampled from everyday mobile devices to calculate personalized estimates of environmental impact and exposure. Crowdsourcing of Pollution Data using Smartphones [144] involves the general public and uses off-the-shelf smartphones as noise sensors. The authors seek to provide a low cost solution for citizens to measure their personal exposure to noise in their everyday environment and participate in the creation of collective noise maps by sharing their geo-localized and annotated measurements with the community.

Air Monitoring Commonsense project [81] exploits participatory sensing systems that allow individuals to measure their personal exposure, groups to aggregate their members' exposure, and activists to mobilize grassroots community action. It is a distributed air quality monitoring system, which combines handheld environmental air quality devices with a browser-accessible web portal. Hasenfratz et al. [12] present the design, implementation, and evaluation of Gas-Mobile, a small and portable measurement system based on off-the-shelf components and suited to be used by a large number of people.

Noise Monitoring NoiseMap [24] is an application that gathers audio samples and transfers them to an open platform to create a real-time map with values of noise. It is evaluated in Frankfurt. Ear-Phone [107] is an end-to-end urban noise mapping system that leverages compressive sensing to solve the problem of recovering the noise map from incomplete and random samples. Delay tolerant networks and WiFi technology are utilized for data delivery. NoiseTube [101] is a noise monitoring platform that exploits GPS and microphone to measure the personal exposure of citizen to noise in everyday life. The main feature is to measure the level of noise, with the possibility to tag a particular type of noise and label a location, also when participants are indoor and gps is not working.

Health Care

HealthAware [75] is a system that exploits accelerometer to monitor daily physical activity and camera to take picture of food. SPA [117] is a smartphone assisted chronic illness self-management system, which facilitates patient involvement exploiting regular feed-back of relevant health data. Dietsense [20] is a system that automatically takes media documentation of dietary choices with just-in-time annotation, efficient review of captured media by participants and easy authoring dissemination of the automatic data collection protocols. Mobile-phone based Patient Compliance System (MPCS) [118] aims to reduce the time-consuming and error-prone processes of existing self-regulation practice to facilitate self-reporting, non-compliance detection, and compliance reminders. The novelty of this work is to apply social behavior theories to engineer the MPCS to positively influence patients' compliance behaviors.

Intelligent Transportation Systems

In [121], the authors exploit social networks to obtain direct feedbacks and potentially very valuable information from people to acquire awareness in ITS. In particular, it verifies the reliability of pollution related social networks feedbacks into ITS systems. ParkNet is a solution for park monitoring and unlike traffic and pothole detection does not require high accuracy in localization. It builds a real-time map of parking availability and provide it to the users that are searching for a parking slot. In [108], Zhou and Li present a system that predicts bus arrival times relying on bus passengers' participatory sensing. The proposed model is based only on the collaborative effort of the participating users and it uses cell tower signals to locate the users, preventing them from battery consumption. Furthermore, it uses accelerometer and microphone to detect when a user is on a bus. Nericell [76] is a system used to monitor road and traffic conditions, which uses different sensors for a rich sensing and detects potholes, bumps, braking and honking. It exploits the piggybacking mechanisms on users' smartphones. The Pothole Patrol [7] detects and reports the surface conditions of roads. It is evaluated from thousands of kilometers of taxi drives in Boston and provides a classification of identified potholes and surface abnormalities. VTrack [120] is a system which estimates travel time challenging with energy consumption and inaccurate position samples. It exploits a HMM (Hidden Markov Model)-based map matching scheme and travel time estimation method that interpolates sparse data to identify the most probable road segments driven by the user and to attribute travel time to those segments. WreckWatch [119] is a formal model that automatically detects traffic accidents using accelerometer and acoustic data. It immediately sends a notification to a central emergency dispatch server and provide photographs, GPS coordinates and data recordings of the situation.

E-Commerce

Mobishop [114] is a distributed computing system designed to collect, process and deliver product price informations from street-side shops to potential buyers. It exploits the camera to scan receipts. In [109] the authors present a participatory sensing paradigm which can be employed to track price dispersion in homogeneous consumer goods even in offline markets. Similar examples are presented in LiveCompare [115], which exploits the bar codes of products and in PetrolWatch [19], which collects automatically fuel prices from gas stations.

Public Safety

ISafe [110] is a system for evaluating the safety of citizens. It exploits data collected from geosocial networks that are related with crimes and data census from Miami county.

Human Behaviour

AndWellness [111] is a personal data collection system that exploits smartphone sensors to collect and analyze data from user experiences. Darwin phones [112] combine collaborative sensing and classification techniques to correlate human behavior and context on smartphones. Darwin is a collaborative reasoning framework based on classifier evolution, model pooling and collaborative inference.

Mobile Sociale Networks

Miluzzo et al. presents the CenceMe application [122], which combines the possibility to use sensor embedded in mobile phones with sharing of sensed and personal information through social networking applications. It takes a user status in terms of the activity, context or habits and shares them in social networks. Micro-Blog [124] is a system in which new kinds of applications-driven challenges are proposed and are compared to the context of the users exploiting also the location. EmotionSense [123] is a platform for social psychological studies based on smartphones and its key idea is to map not only activities but also emotions and to understand the correlation between them. It gathers users' emotions as well as proximity and patterns of conversation by processing the audio from the microphone. MobiClique [22] leverages already existing social networks and opportunistic contacts between mobile devices to create ad hoc communities for social networking and social graph based opportunistic communications. MIT's Serendipity [127] is one of the first projects that explored the aspects of mobile social networking. It is used to build informal interactions using the combination of bluetooth hardware. SociableSense [125] is a platform that realizes an adaptive sampling scheme based on learning methods and a dynamic computation distribution mechanism based on decision theory. This system captures user behaviour in office environments and provides the participants a quantitative measure of sociability of them and their colleagues. WhozThat [128] presents an opportunistic connectivity for offering an entire ecosystem on which increasingly complex context-aware applications can be built. MoVi [126], a Mobile phone based Video highlights system, is a collaborative information distillation tool capable of filtering events of social relevance. It consists on a trigger detection module that analyzes the sensed data of several social groups and recognize potentially interesting events

Others

Crowdsense@place [145] provides place related informations, including relationship between user and coverage. It exploits a research on scaling properties of place-centric crowdsensing [132] and presents money as incentive scheme. In [129], the authors improve the location reliability, proposing a scheme in which participatory sensing is used to achieve data reliability. The key idea of this system is the location validation using photo tasks and expanding the trust to nearby data points using periodic bluetooth scanning. The participants are asked to send a number of photo tasks from the known location, which are manually or automatically validated. ILR is evaluated in the context of McSense. It presents amazon mturk [146] as incentive. SoundSense [73] is a scalable framework for modelling sound events on smartphones, using a combination of supervised and unsupervised learning techniques to classify both general sound types and discover sound events specific to individual users. Travel packages are proposed in [137], exploiting a recommendation system to help users in planning travels by leveraging data collected from crowdsensing. The authors propose to distinguish user preferences, extract points of interest (POI) and determine location correlations from data collected. Consequently, personalized travel packages are determined by considering personal preferences, POI, temporal and spatial correlations. In ConferenceSense [130] collected data is used to extract and understand community properties to sense large events like conferences. It uses some sensors and user inputs to infer contexts such as the beginning and the end of a group activity. Guo et al. [147] propose a model for sensing the volume of wireless activity at any frequency exploiting the passive interference power. This technique utilizes a non-intrusive way of inferring the level of wireless traffic, without extracting data from devices. Furthermore, the presented approach is independent from the traffic pattern and requires only approximate location data. In [58], Farshad et al. exploit MCS for urban WiFi characterization and monitoring, measuring spectrum and interference in the city of Edinburgh. MCNet [131] enables WiFi performance measurements taken from users that participate in the sensing process. The authors present a system for collecting and mapping data about WLAN with a crowdsensing approach. SorroundSense [148] is a system that explores logical localization, exploiting ambience fingerprinting. The key idea is that the combination of ambient sound, light and colour can be unique enough for localization and to distinguish a place from another one.

Table 3.2: Classification on Types of Sensors

Target	Type of Sensors										
	Acc	Gyr	Mag	GPS	Mic	Cam	Tem	Hum	WiFi	Blt	Cts	NoE
EC				[114][19] [115]		[114][19] [115]			[115]		[115]	
EM	[116][12]	[116] [12]	[116] [12]	[12][81] [106][116] [24][107] [101]	[24][107] [101]		[12]	[12]	[24]		[106][24]	[12][81]
HC	[75][117] [118]	[75]		[20][75] [117][118]	[20][117]	[20][75]	[117]		[117]	[117]		[117]
HB				[111][112]	[112]	[112]						[112]
ITS	[76][7] [108][119]	[108] [119]	[108] [119]	[76][7] [120][121] [9][119]	[76][108] [119]	[76][121] [119]						[121][9]
MSN	[122][123] [124][125] [126]	[125] [126]	[126]	[122][124]	[122][123] [124][125] [126]	[122][124] [126]				[122][22] [127][125] [128]		
Oth	[129][130] [131]	[130] [131]	[130] [131]	[132][129] [58]	[132][130] [73]	[132][129]			[58][131]	[129][130]		[129][130]

Table 3.3: Classification on Sensing Activity

Target	Sensing Activity									
	User Recruitment		User Involvement		Decision Process		Scope		Reporting	
	Volunteer	Incentive	Participatory	Opportunistic	Mobile device	Collector	Individual	Community	Infrastructure	Ad-hoc
EC	[19][115]	[114]	[19][114] [115]		[19][114] [115]			[19][114] [115]	[19][114] [115]	[114][115]
EM	[12][81] [106][116]	[24][107]	[12][81] [106][107]	[116][24]	[106][24] [107]	[12][81] [116]		[12][81] [106][116] [24][107]	[106][116] [24]	[116][107]
HC	[75][117]		[75][117]		[75][117]		[75][117]		[75]	[75][117]
HB	[111][112]		[111][112]		[111][112]		[111]	[112]	[112]	[111][112]
ITS	[121][9] [108]	[76][7] [120][119]	[121]	[76][7] [120][9] [108][119]	[76][7] [120][9] [108][119]	[121]	[120][108] [119]	[76][7] [121][9]	[76][7] [121] [108][119]	[76][7] [120] [121][9]
MSN	[122][124] [22][127] [123]	[125][128] [126]	[122][124] [22][127] [125][128] [126]	[123]	[122][123] [124][22] [127][125] [128]	[126]		[122][123] [124][22] [127][125] [128][126]	[124][126]	[122][123] [124][22] [127][125] [128]
OTH	[130][137] [73][132]	[129][58] [131]	[129][130] [137][73]	[132][58] [131]	[132][129] [130][137] [58][73] [131]		[73][58] [131]	[132][129] [130][137]	[137][73]	[132][129] [130][137] [58][73] [131]

Table 3.4: Classification on Sensing Task

Target	Sensing Task									
	Assignment		Distribution		Execution Time		Execution Method		Frequency	
	A-priori	A-posteriori	Centralized	Decentralized	Synchronous	Asynchronous	Independent	Collaborative	Event	Continuous
EC		[114][19] [115]	[114][19] [115]		[115]	[114][19]		[114] [19] [115]	[114] [19] [115]	
EM	[12][81] [24][107] [101]	[116][106]	[12][81] [106][116] [24][107] [101]		[116]	[12][81] [106][116] [24][107] [101]	[12][81] [106][116] [24][107] [101]		[81]	[12][81] [106][116] [24][107] [101]
HC	[20]	[75][117] [118]	[20]	[75][117] [118]		[20][75] [117][118]	[20][75] [117]	[118]	[20][75]	[117][118]
HB		[111][112]	[111]	[112]		[111][112]				
ITS	[9][108]	[121]	[9]	[121][108]		[121][9][108]	[121][9]	[108]	[121][108]	[9]
MSN		[122][123] [124][125] [126][22] [127][128]	[122][123] [124][125] [126]	[22][127] [128]		[122][123] [124][125] [126][22] [127][128]	[122][123] [125][127]	[124][126] [22][128]	[124][126]	[122][123] [125][22] [127][128]
Oth	[58][73] [131]	[132][129] [130][137]	[130][137] [58][73] [131]	[132][129]		[132][129] [130][137] [58][73] [131]	[132][130] [137][58] [73][131]	[129]	[132][129] [130][137]	[73][58] [131]

Table 3.5: Description of Properties of Data

Target	Properties of Data							
	Type				Delivery		Granularity	
	Value	Image	Audio	Position	Real Time	Delay Tolerant	Fine-grained	Coarse-grained
EC	[114]	[19][115]		[114][19][115]	[115]	[114][19]	[114][19][115]	
EM	[12][81][116] [106][24] [107][101]			[12][81] [106][24] [107][101]	[106][116] [24][101]	[12][81] [116][107]	[12][81] [116][24] [107][101]	[106]
HC	[75][117][118]	[20][75]	[20][117]	[20][75] [117][118]	[20][117]	[20][75][118]	[20][75][117]	[118]
HB	[112]	[112]	[112]	[111][112]		[111][112]	[112]	[111][112]
ITS	[121][9][108]	[121]		[121][9][108]	[9][108]	[121]		[121][9][108]
MSN	[122][123] [124][125] [126][22] [127][128]	[122][124] [126]	[122][123] [124][125] [126]	[122][124]	[122][124] [125][128]	[123][126] [22][127]	[122][123] [124][125]	[126][22] [127][128]
Oth	[132][130] [137][58] [131]	[132][129]	[132][130]	[132][129] [137][58] [131]		[132][129] [130][137] [58][131]	[132][129] [137][58] [131]	[130]

Chapter 4

Data Collection Framework for Mobile Crowd Sensing Systems

Existing studies in MCS systems focus on developing efficient incentive mechanisms to foster user participation, while data collection algorithms still require investigation efforts. As illustrated in Sec. 3.2.3, there are two MCS paradigms that involve users in the data collection process: participatory and opportunistic [82, 83]. In *opportunistic* sensing, the user involvement is minimal or null, which means that the decisions to perform sensing and report data are application- or device-driven. On the other hand, in *participatory* sensing the user is actively engaged in the process. For example, a participant is in charge of capturing an image or to record a video by deciding whether, when and for how long sensing is performed. Participatory sensing is tailored to crowd sensing architectures with a "central intelligence" responsible to task users, e.g. to ask one user to record a video in a given area at a given time. This imposes a higher cost to the user than opportunistic sensing in terms of cooperation effort. Having devices or applications responsible for sensing decision, it lowers the burden for users participation and makes opportunistic sensing ideal for distributed solutions.

This chapter proposes a novel distributed algorithm for gathering information in cloud-based MCS systems, exploiting the opportunistic sensing paradigm. This framework minimizes the cost of both sensing and reporting operations while maximizing at the same time the utility of data collection, i.e. the quality of contributed information. Fig. 3.1 illustrates the reference scenario considered. The crowd exploits mobile devices sensors to contribute information, which is delivered to a collector in the cloud that is responsible for data processing and analysis. In the rest of the chapter, the terms crowd, participants and users will be used interchangeably.

The data collection algorithm is cost-effective for the participants, minimizing the cost they experience in performing sensing and reporting. The cost is measured in terms of the energy spent by the devices. At the same time, the proposed algorithm maximizes the utility of samples the collector receives. To this end, the collector broadcasts beacons to the participants to inform them about the most urgent samples to collect. This is defined as the *data collection utility*, which can also be seen as the quality of contributed information for the system guaranteed by the algorithm. Indeed, the cloud collector broadcasts periodically messages to the users indicating which type and the amount of data required to capture a physical phenomena while ensuring a minimum level of accuracy in the process.

The participants determine whether they can contribute such data on the basis of their *sensing potential*, which is the cost of sensing and reporting. To maximize the utility of data collection, it is fundamental to take into account the *environmental context* of the devices [149, 150]. For example, capturing a picture while the smartphone is in a pocket does not bring any utility for the MCS system. Users sense and report data only when there is match between the smartphone sensing potential and the data collection utility, as shown in Fig. 4.1. The following sections describe the data collection policies that overcome such issues and validate the proposed algorithm both analytically and through simulations. Table 4.1 lists a description of symbols used in the following sections.

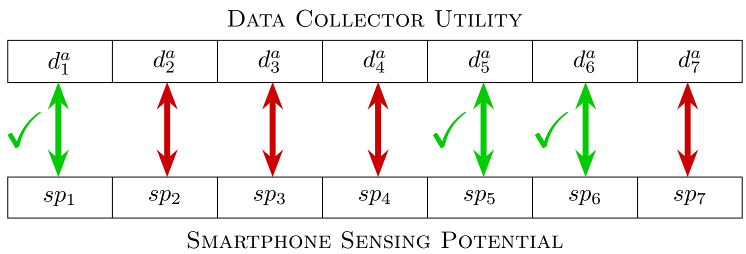


Figure 4.1: Matching between the smartphone sensing potential and the data collection utility

4.1 Data Collection Policies

The mobile devices decide to perform sensing and reporting operations independently one each other on the basis of the *data collection utility* d_s^a, the *smartphone sensing potential* sp_s and the *environmental context* of the device C_s. The latter parameters are determined locally at the mobile devices while the former is computed by the cloud collector. The objectives of the entities, the cloud collector and the crowd, may be in contrast. For example, the collector may require samples from a given sensor while all the participants in that area may want to preserve their resources.

Table 4.1: Symbols list and description

Symbol	Description
s	Sensor s
$\mathcal{S}$	Set of sensors
a	Monitored area
$\mathcal{A}$	Set of monitored areas
t	Timeslot
d_s^a	Data collection utility for sensor s in area a
sp_s	Smartphone sensing potential for sensor s
C_s	Environmental context for sensor s
$N_s^a\|_t$	Number of samples generated by sensor s in timeslot t in area a
E_s	Energy of sensor s
E_s^c	Energy spent for sensing of sensor s
U_s	Utilization context for sensor s
E_s^r	Energy spent by sensor s for reporting
P^W	Power consumption of WiFi
P^L	Power consumption of LTE
l_s^x	Last sensed value
l_s^r	Last reported value
ϵ_s	Threshold for deciding if report last sample
B	Level of Battery
D	Amount of the reported data
D_s	Amount of the reported data from sensor s
D_s^W	Amount of the reported data via WiFi
D_s^L	Amount of the reported data via LTE
δ	Threshold for making sensing and reporting decisions
δ_b	Contribution of the level of battery to δ
δ_d	Contribution of the amount of reported data to δ

To compare the previous parameters with a threshold δ, two different data collection policies are defined: a *collector-friendly* policy and a *smartphone-friendly* one. The former takes into consideration the *data collection utility* with higher priority while in the second policy the costs for sensing and reporting assume more relevance.

Collector-friendly Policy (CFP) The policy is defined as follows:

$$[C_s \cdot \gamma \cdot sp_s] + (1-\gamma) \cdot d_s^a > \delta, \tag{4.1}$$

where γ is a coefficient that can assume real values in the range [0,1]. High values of γ give more relevance to the *smartphone sensing potential*, while low values of γ make dominant the *data collection utility* term.

Smartphone-friendly Policy (SFP) The policy is defined as follows:

$$C_s \cdot [\gamma \cdot sp_s + (1-\gamma) \cdot d_s^a] > \delta. \tag{4.2}$$

With respect to the *collector-friendly* policy defined in (4.1), when the environmental context is unfavorable, i.e. $C_s = 0$, the devices never perform sensing and reporting. Indeed, under such hypothesis the utility of the samples would be very low for the collector. As a result, the policy prevents the devices to perform useless operations to save resources. On the other hand, the adoption of the *collector-friendly* policy would make the devices to perform sensing and reporting even if the environmental context is unfavorable. This entails the collector receiving much more data and can be useful to monitor phenomena that require many samples to be captured or when little amount of information comes from a given area.

4.2 Data Collection Utility at Cloud Collector

Following the Sensing-as-a-Service (S^{2}aaS) paradigm, the data collector is located in the cloud (see Fig 3.1). S^{2}aaS paradigm has recently gained significant attention from the research community [151, 152, 153]. Initially, the users perform a request to the cloud for sensing data in a given area. The sensors and the mobile devices located in that area accomplish the request and collect the information. Then, the information is delivered to the cloud, stored and made available to the cloud users. Such model perfectly suits MCS and it is of special interest when applied to smart cities [154].

On the basis of the requests for sensing that are assumed to be known in advance, the cloud collector decides which samples need to be collected. Each request can come from different applications and requires samples from different location areas and type of sensors. As a result, it is possible to identify a utility function in the data collection process that typically exhibits the so called *marginal effect* [155]. Areas with higher user participation will satisfy quickly the demand of samples leading to low utility in gathering more information. On the other hand, having few samples at disposal makes the need for data urgent and the utility of such data becomes high. However, the marginal effect in data collection applies only in the case a given phenomena has to be captured once. When the objective is to perform continuous monitoring, more samples become at disposal in the cloud collector, the higher the accuracy in mapping the phenomena.

To define the *data collection utility*, the monitored area is assumed to be partitioned in a set of areas $\mathcal{A}$. In each area $a \in \mathcal{A}$, the mobile devices generate samples from a different set of sensors $\mathcal{S}$. The Exponential Weighted Moving Average filter (EWMA) describes the average number of samples $\overline{N}_s^a|_t$ generated from the sensor s in the area a during the timeslot t:

$$\overline{N}_s^a|_t = \sigma \cdot N_s^a|_t + (1-\sigma) \cdot \overline{N}_{s-1}^a|_t, \tag{4.3}$$

where $N_s^a|_t$ corresponds to the number of samples of sensor s in timeslot t in area a and $\overline{N}_s^a|_{t-1}$ is the previous value. The parameter σ is the exponential weighting coefficient. High values of σ limit the contribution of older values.

When $\overline{N}_s^a|_t$ assumes high values, the cloud collector has received many samples. As a consequence, the data collection utility is low. On the other hand, low values of $\overline{N}_s^a|_t$ indicate the need for more samples and the data collection utility is high. These considerations suggest that the data collection utility can be defined as a sigmoid function (shown in Fig. 4.3):

$$d_s^a = \frac{1}{1 + \mathrm{e}^{-\frac{\varphi_s}{\rho_s} \cdot (-\overline{N}_s^a|_t + (1 - \frac{\rho_s}{2}))}}, \tag{4.4}$$

where φ_s and $1 - \rho_s/2$ represent the incline and the center of the curve respectively. The data collection utility can assume real values in the range [0,1].

The cloud collector computes d_s^a per area a and sensor s. Then, it informs the participants in each area a by means of beacon messages transmitted periodically.

4.3 Smartphone Sensing Potential

This section models the potential of the mobile devices in performing sensing. Describing the *smartphone sensing potential* implies two parameters: the energy spent for both sensing and reporting operations and the utilization context. Taking into account both parameters it is fundamental to characterize the utility in collecting data from the user point of view. Understanding the environmental context of the devices involves utilization of always-on sensors such as the accelerometer, which obviously has a cost in terms of energy. However, having such a knowledge helps to avoid performing more costly and not useful sensing operations, such as taking photo when the smartphone is in a pocket.

The energy the mobile devices spend to contribute data can be attributed to sensing and reporting operations. For each sensor s, the energy cost E_s is therefore defined as:

$$E_s = E_s^c + E_s^r. \tag{4.5}$$

For sensing, the contribution E_s^c has to be taken into account only if the sensor s is not already in use by another application. In such a case, similarly to [156], the energy spent by sensor s is equal to zero. This behavior is called utilization context of s, U_s. As a result, E_s^c is defined as follows:

$$E_s^c = \overline{E}_s^c \cdot U_s, \tag{4.6}$$

where $\overline{E}_s^c$ is the actual energy spent by sensor s and the utilization context U_s is defined as:

$$U_s = \begin{cases} 0, & \text{if the sensor } s \text{ is used by another application;} \\ 1, & \text{otherwise.} \end{cases} \tag{4.7}$$

Data reporting consists in delivering the information gathered from the set sensor $\mathcal{S}$ to the cloud collector. Reporting is assumed to take place always at the beginning of the timeslot t for the samples collected during timeslot $t-1$. The contribution E_s^r to the energy cost E_s depends on the technology used for communication with the cloud. Cellular (LTE) and WiFi technologies are both taken into account. E_s^r is defined as follows:

$$E_s^r = \begin{cases} E^W, & \text{if WiFi or both WiFi and LTE are on;} \\ E^{L_1}, & \text{if WiFi is off and LTE is idle state;} \\ E^{L_2}, & \text{if WiFi is off LTE is connected state.} \end{cases} \tag{4.8}$$

Most of the operating systems for smartphones, including Android and iOS, already make preference to WiFi over cellular connectivity for data transmission when both are available. Indeed, WiFi is much more energy efficient if compared to LTE [157] and users do not consume the data plan they pay to the cellular operators [158]. As a result, when both WiFi and LTE interfaces are active, WiFi is preferred. To model the power consumption for data transmission with WiFi P^W, it is used the model proposed in [159]:

$$P^W = \rho_{\text{id}} + P_{\text{tx}}^W + P_{\text{xg}}(\lambda_g), \tag{4.9}$$

The energy E^W spent during the transmission time T_{tx} is defined as:

$$E^W = \int_0^{T_{\text{tx}}} P_{\text{tx}}^W \, \mathrm{d}t. \tag{4.10}$$

On the other hand, to model the energy consumption of LTE it is taken into account the RRC state machine and the simplified model proposed in [160], in which there are different energy terms related to the state of the machine. If the User Equipment (UE) is in idle state and it sends data, before transmitting there is a promotion state from *idle* to *connected* and after the transmission there is a tail before coming back to idle. The energy consumption for the smartphone due to the reporting is:

$$E^{L_1} = P_\mathrm{P} \cdot T_\mathrm{P} + P^L_\mathrm{tx} \cdot T_\mathrm{tx} + P^L_\mathrm{tx} \cdot DRX_\mathrm{IT} + P_\mathrm{DRX} \cdot RRC_\mathrm{IT}, \tag{4.11}$$

where T_P and P_P are the promotion delay and power, T_tx and P_tx are time and power transmission, DRX_IT is the DRX Inactivity Timer, P_DRX is the power consumed when the UE is in one of the two DRX modes and RRC_IT is the RRC Inactivity Timer. Otherwise, if the smartphone is already connected to LTE and transmitting, the energy consumption is given only by the transmission contribute:

$$E^{L_2} = \int_0^{T_\mathrm{tx}} P^L_\mathrm{tx} \, \mathrm{d}t. \tag{4.12}$$

The power consumption for transmitted data P_tx is given by the model of [157]:

$$P^L_\mathrm{tx} = \alpha_\mathrm{ul} \cdot T_\mathrm{ul} + \beta, \tag{4.13}$$

where T_ul represents the uplink throughput and the parameters α_ul and β can be found in the linear model proposed in [157].

In order to maximize energy savings and data collection utility at the cloud collector, reporting is not performed in case the value of the last sensed sample l^x_s is similar to the last reported value l^r_s for sensor s. Specifically:

$$|l^x_s - l^r_s| < \epsilon, \tag{4.14}$$

where ϵ is a threshold defining the similarity between the two values.

This operation can provide significant energy savings especially for sensors generating samples with values that do not exhibit high variability. For example, the sensors used for meteorological measurements like the temperature and humidity sensor and the barometer. When the variation between the values of samples generated from the same sensor s in subsequent timeslots remains little, the utility in reporting the last sample is low. As a consequence, to maximize the utility at the cloud collector and save energy, the algorithms prevents reporting and stops the sensing process. This procedure is modeled defining the timeslot in which the next sample will be collected as:

$$t_\mathrm{next} = \begin{cases} t + i \cdot n, & \text{if } |l^x_s - l^r_s| < \epsilon; \\ t + 1, & \text{otherwise.} \end{cases} \tag{4.15}$$

The parameter n is a fixed number of timeslot to skip and i is the number of times the variation between the values of the samples remains below ϵ. In case the

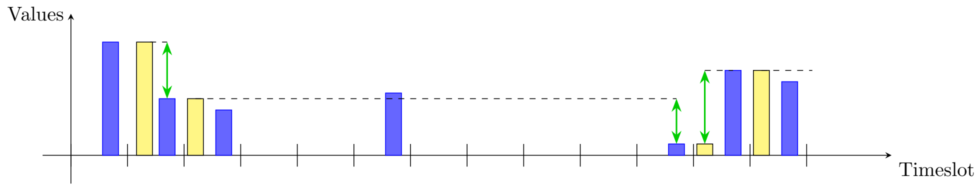


Figure 4.2: Example for sensed and reported samples

decision is to not report, the cost of reporting E_s^r assumes values equal to zero. An example is shown in Fig.4.2.

The smartphone sensing potential sp_s for each sensor s is function of the energy cost E_s. Similarly to the data collection utility, the relation is described as a sigmoid function (shown in Fig. 4.3):

$$sp_s = \frac{1}{1 + e^{-\frac{\zeta_s}{\theta_s} \cdot (-E_s + (1 - \frac{\theta_s}{2}))}}, \tag{4.16}$$

where ζ_s and $1 - \theta_s/2$ represent the incline and the center of the curve respectively. The smartphone sensing potential can assume real values in the range [0,1].

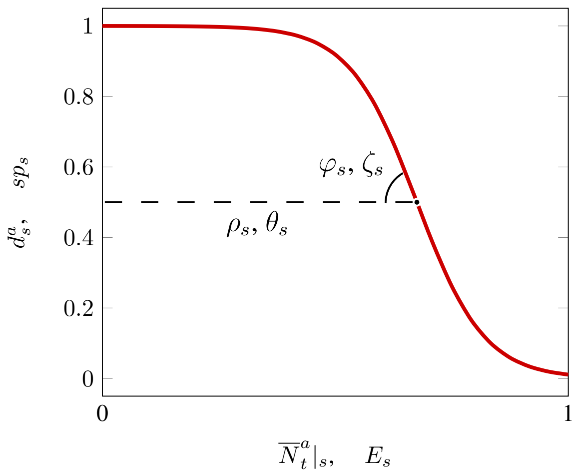


Figure 4.3: Data Collection Utility and Smartphone Sensing Potential

4.4 Profiling the Environmental Context

The *environmental context* C_s of sensor s is defined as the set of facts and circumstances happening around the mobile device. Note that such a definition not only takes into account the location of the mobile device (e.g., lying on a table, in a pocket, in a hand, indoor or outdoor), but also the relation with the user movement (e.g., stationary or on the move, with the possibility to recognize the type of movement thanks to several movement pattern). Context awareness is typically performed with always-on sensors, such as the accelerometer that enables recognition of movement patterns (e.g. walking or running [25]) and actions (e.g. driving, riding a car or sitting [27]). Estimating the environmental context is not simple, although it exists practical solutions [161]. However, its knowledge is crucial to avoid performing useless sensing operations that impact on the overall energy budget without providing any benefit for the collector, such as taking photo when the smartphone is in a pocket.

The utility of each sample is defined according to the environmental context. For the sake of simplicity and without loss of generality, C_s assumes binary values:

$$C_s = \begin{cases} 1 & \text{if the sample brings some utility in the process;} \\ 0 & \text{otherwise.} \end{cases} \tag{4.17}$$

In the proposed framework the collector builds a table for each sensor at the beginning of the collecting process. The table is built according to the contexts in which the collector is interested in and it is sent to the smartphones in each monitored area at the beginning of the sensing process. For instance, if the data collector needs the position of a device, the GPS will be useful only outdoor. It follows that C_s will be set to 1 for outdoor environments and set to 0 for indoor environments. Consequently, when smartphones receive the table at the beginning of the sensing process, they know case by case how to set the bit corresponding to C_s according to the environmental context they reveal. Each time the environmental context of the mobile devices changes, the smartphones sense it and assign a new binary value to C_s. The table containing context profiles the collector broadcasts to the users can be updated anytime. For example, this happens when the collector changes sensing task, e.g. from air monitoring to gluten detection.

4.5 Threshold for Sensing and Reporting

Sensing and reporting operations occur when the parameters *data collection utility* and *smartphone sensing potential* are greater than a threshold δ. This translates in having the mobile devices sustaining a cost that produces useful data for the cloud collector. In addition, this mechanism can be tuned to prevent users to contribute too much or too little. The threshold δ becomes therefore a key parameter and needs to be set properly. To this end, it is considered the actual *level of the battery* of the devices B and the *amount of reported data* D the devices have already contributed to the system. The two parameter define the historical cost the devices have sustained. The parameter δ is defined as follows:

$$\delta = \begin{cases} 1, & \text{if } \delta_d = 1 || \delta_b = 1; \\ f(\delta_B, \delta_d), & \text{otherwise;} \end{cases} \tag{4.18}$$

where all the parameters δ, δ_b and δ_d are real values in the range [0,1]. Specifically, both δ_b and δ_d are function of B and D respectively. Having any of the two parameters assuming values equal to 1 makes the threshold δ to become 1 as well. As a result, the devices will stop contributing to the system. The function f can be chosen arbitrarily. In this thesis, for simplicity it is assumed to provide equal weight to δ_b and δ_d:

$$f(\delta_b, \delta_d) = \frac{\delta_b + \delta_d}{2}. \tag{4.19}$$

Level of Battery The *level of battery B* is remaining charge of the device, where $0 \le B \le 100$. High values of B denotes a high level of charge and the device can contribute data from its sensors. On the other hand, when the battery is almost empty the users would like to preserve the remaining resources. These considerations suggests that the relation between B and δ_b follows a negative exponential law. As a result, low values of B will make δ_b to assume values close to 1. In addition, B_{min} is defined as the minimum level of the battery under which the device stops contributing. Note that $0 \le B_{\text{min}} < B$. As a consequence, δ_b is defined as follows (as shown in Fig. 4.4):

$$\delta_B = \begin{cases} \alpha^{\lambda \cdot b}, & B_{\text{min}} \leqslant b < 100; \\ 0, & \text{otherwise.} \end{cases} \tag{4.20}$$

The parameter α can assume arbitrary real values between [0,1] and λ is always greater than 1. Proper setting of the parameter B_{min} plays a crucial role. Values of B_{min} close to B lead the device to contribute little amount of data. On the other hand, it will make the device to contribute more.

Amount of Reported Data The *amount of reported data D* is total amount of data contributed by the set of sensors $\mathcal{S}$ of a single device:

$$D = \sum_{s \in \mathcal{S}} D_s. \tag{4.21}$$

Specifically, the contribution of data delivered using WiFi (D_s^W) and LTE (D_s^L) is considered:

$$D_s = D_s^W + D_s^L. \tag{4.22}$$

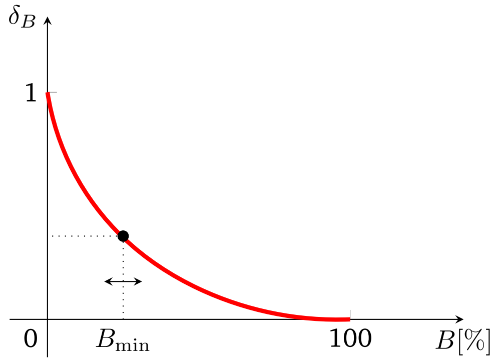


Figure 4.4: Contribution δ_b of the level of battery to δ

The more data the device contributes, the higher the values parameter δ_d assumes. As a result, the devices that have already delivered a significant amount of data to the cloud collector in the past contribute further only if it is necessary. On the other hand, low values of the parameter δ_d will enforce the contribution of the devices that have provided little contribution to the system. As a result, the relation between D and δ_d is modeled as follows (as shown in Fig. 4.5):

$$\delta_d = \begin{cases} 1, & \text{if } D \geq D_{\max}; \\ \log\left(1 + \dfrac{D}{D_{\max}}\right), & \text{otherwise}; \end{cases} \tag{4.23}$$

where $D_{\max}$ is the maximum amount of data each device can deliver. This parameter can be tuned by the users periodically.

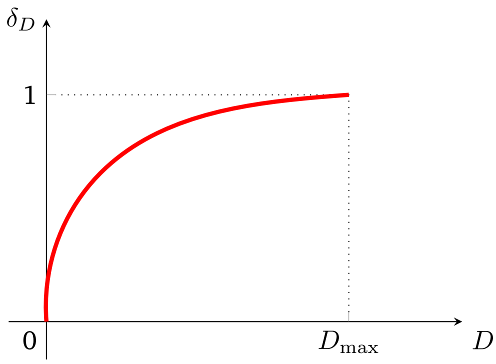


Figure 4.5: Contribution δ_d of the amount of reported data to δ

Chapter 5

Performance Evaluation

This chapter illustrates the performance evaluation of the proposed framework by providing an analytical study. The following sections present the set-up of the evaluation and the results, which assess the amount of data generated by the mobile devices and the energy consumed for sensing and reporting.

5.1 Evaluation Set-up

For the evaluation, four users are considered and each of them is equipped with one device only. The devices have different level of remaining battery charge: 80%, 100%, 10% and 50%. The battery charge is one of the parameters affecting the threshold δ for sensing and reporting (parameter δ_b). Recalling equation (4.20), to compute δ_b it is set $\alpha = 0.7$ and $\lambda = 10$. The other parameter, δ_d, corresponds to the amount of previous reported data and is set to 0 for all the users at the beginning of the evaluation. The evaluation period consists of 600 timeslots and each timeslot corresponds to 1 s.

Each mobile device generates data from the set of sensors $\mathcal{S}$, which includes the accelerometer, the temperature and the pressure sensor and transmits the information over WiFi. Without loss of generality, having only one communication technology makes easy the understanding of the properties of the data collection framework. For the sensing equipment, real sensors implemented in current smartphones and tablets are taken into account. Specifically, the FXOS8700CQ 3-axis linear accelerometer from Freescale Semiconductor and the BMP280 from Bosch, which is a digital pressure and temperature sensor. Equation (4.9) describes WiFi power consumption spent by the devices for communication. Table 5.1 presents the detailed information on communication and the parameters.

For the sake of simplicity, all the users obtain the same profiles for the environmental context from the collector: $C = \{1.0, 1.0, 1.0\}$. For each timeslot, each user

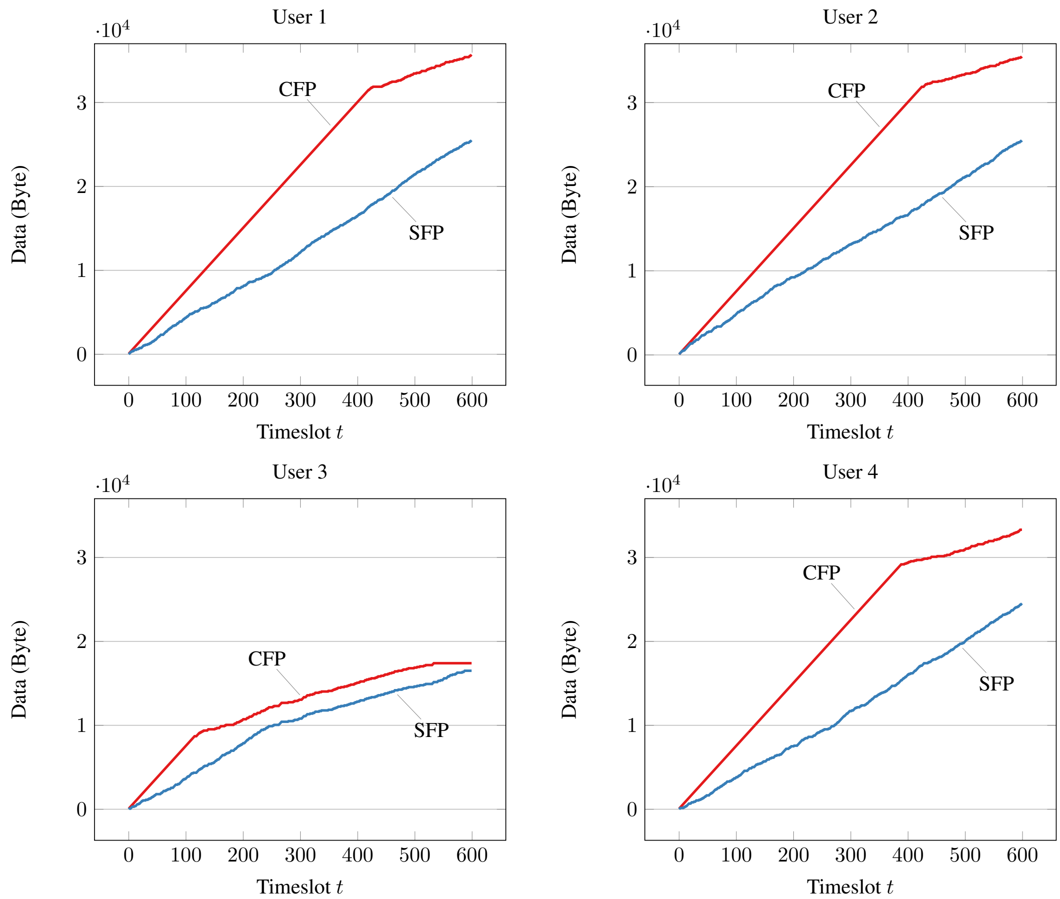


Figure 5.1: Amount of data collected per user from accelerometer readings

compares the current context with the profiles. The current context c_t is a random variable generated in each timeslot t and it follows an exponential distribution with rate 1.25. According to C, values of c_t close to 1.0 correspond in having the smartphone performing sensing in favorable conditions. On the other hand, values of c_t close to 0.0 correspond to samples generated in unfavorable conditions and therefore such data does not bring any utility from the collector point of view. The utilization context U_s is modeled as a random variable that assumes values uniformly distributed in the range [0,1] and is generated during each timeslot for each user device.

Table 5.1: Sensor and communication equipment parameters used for performance evaluation

(a) Sensor

SENSOR	PARAMETER	VALUE	UNIT
Accelerometer	Sample rate	50	Hz
	Sample size	12	Bits
	Current	35	µA
Temperature	Sample rate	182	Hz
	Sample size	16	Bits
	Current	182	µA
Pressure	Sample rate	157	Hz
	Sample size	16	Bits
	Current	423.9	µA

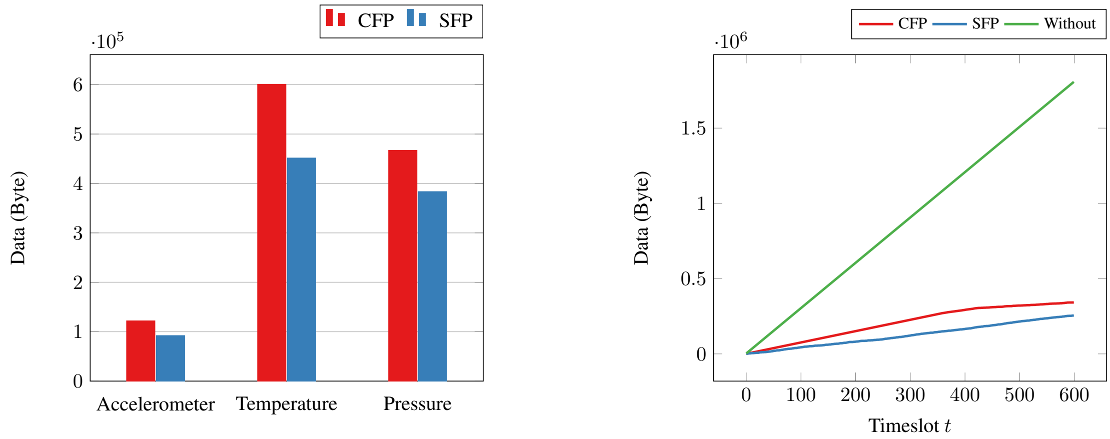
(b) Communication

SYMBOL	VALUE	UNIT	DESCRIPTION
ρ_{id}	3.68	W	Energy in idle mode
ρ_{tx}	0.37	W	Transmission power
ρ_{rx}	0.31	W	Reception power
λ_r	1000	fps	Rate of received packets
λ_g	1000	fps	Rate of generation of packets
γ_{xr}	$0.09 \cdot 10^{-3}$	J	Energy to elaborate a received packet
γ_{xg}	$0.11 \cdot 10^{-3}$	J	Energy to elaborate a generated packet

5.2 Results

The objective of the evaluation is in assessing the amount of data generated and the energy consumption experienced for performing sensing and reporting. In addition, this analysis studies the overall performance of the framework. To this end, it is evaluated the impact of the parameters such as δ, the threshold for sensing and reporting and γ the coefficient balancing the impact of the data collection utility and the smartphone sensing potential.

Having set $\gamma = 0.5$, Fig. 5.1 illustrates the amount of data the four users generates from accelerometer readings. The other sensors exhibit a similar behavior and for space reasons the results are not shown. The performance of both CFP and SFP policies are investigated for the entire evaluation period. As it is possible to notice, the CFP policy makes the users to contribute higher amounts data than the SFP policy. This behavior is expected as the SFP is tailored to perform operations



(a) Analysis per sensor

(b) Comparison of policies

Figure 5.2: Total amount of collected data

only when the environmental context guarantees the devices to generate useful information to the system. It is interesting to notice that User 3 contributes nearly the same amount of data with both policies. Being the remaining charge at the beginning of the simulation low, the threshold δ assumes values close to one. Having set $\gamma = 0.5$ the data collection utility and the smartphone sensing potential accounts in the same manner in order to take sensing and reporting decisions. As a result, User 3 provides a contribution to the system only when the collector has a high interest in obtaining samples from the user. For all the users but User 3, the amount of contributed data has a sharp increase that becomes smoother towards the end of the evaluation period. The motivation is twofold. On one side, having obtained enough samples the utility at the collector become smaller (data collection utility factor). On the other hand, the remaining capacities of the devices decrease with time while the amount of already reported data increase. As a result it becomes more difficult to meet the threshold δ and this lowers the contribution rate.

Fig. 5.2 analyzes the total amount of collected data. In particular, Fig. 5.2(a) compares the amount of information each sensor generates. Being proportional to the sampling frequency of each sensor, the accelerometer generates the lower amount of data in comparison with the temperature and pressure sensor. Low values of sampling frequency make the difference between the amount of generated data by the to policies to be small. On the other hand, with the increase of the sampling frequency such difference increases. Obviously the amount data generated is proportional to the sampling rate and resolution size, being the temperature sensor having the highest rate and resolution size (see Table 5.1(a)).

Fig. 5.2(b) provides a comparison between the total amount of data generated by the two policies and when our framework is not utilized. Each curve denotes the amount of data generated by the three sensors together. As expected the amount of data generated without using the proposed framework is higher than having active either CFP or SFP. This is because the devices perform sensing and reporting even when the environmental context c_t during each timeslot t is unfavorable. As a result, many of the generated samples do not bring any additional utility to the system. The SFP makes the user to only perform operations when the environmental context is favorable and for this reason generates the lowest amount of data. On the other hand, the CFP is more flexible and allows the devices to sense and report data even in case the environmental context is unfavorable, but the data collection utility and the smartphone sensing potential conditions should always be met. As a consequence, the amount of data generated by means of the CFP is significantly lower than the case when the framework is not utilized.

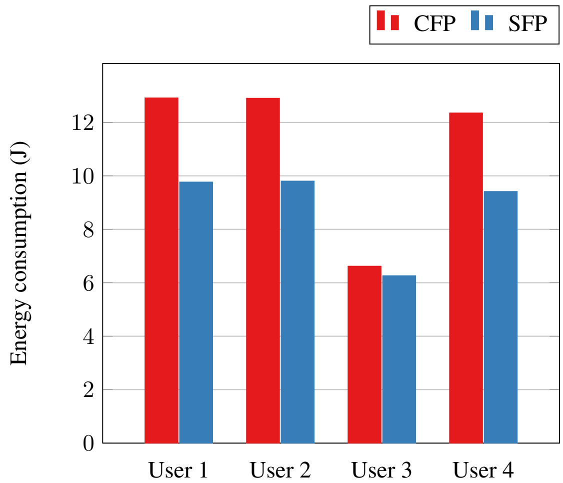


Figure 5.3: Energy consumption for the data collection policies

Fig. 5.4 provides a comparison between the total amount of data generated by the two policies and when the framework is not utilized. Each curve denotes the amount of data generated by the three sensors together. As expected the amount of data generated without using the proposed framework is higher than having active either CFP or SFP. This is because the devices perform sensing and reporting even when the surrounding context c_t during each timeslot t is unfavorable. As a result, many of the generated samples do not bring any additional utility to the system. The SFP makes the user to only perform operations when the surrounding context is favorable and for this reason generates the lowest amount of data. On the other hand, the CFP is more flexible and allows the devices to sense and report data even in case the surrounding context is unfavorable, but the data collection utility and the smartphone sensing potential conditions should always be met. As a consequence, the amount of data generated by means of the CFP is significantly lower than the case when the framework is not utilized.

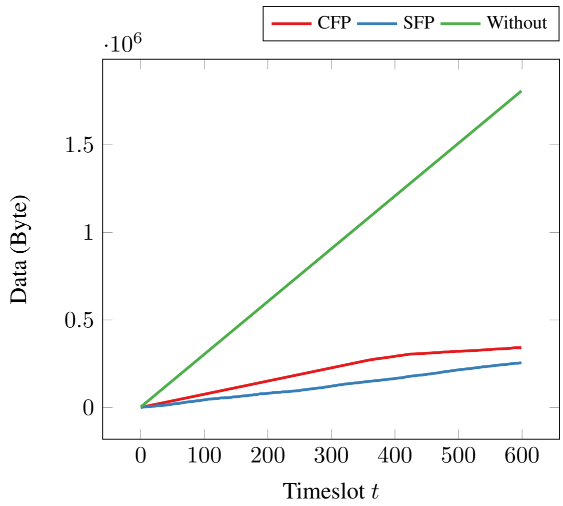
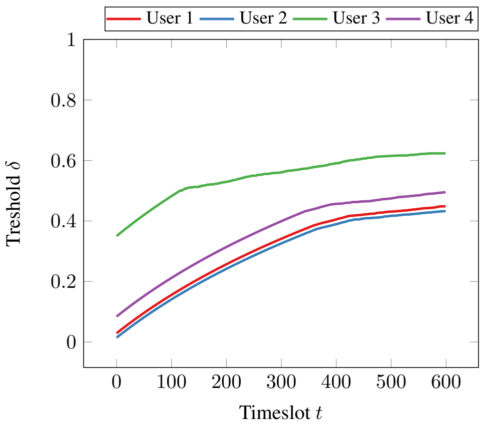


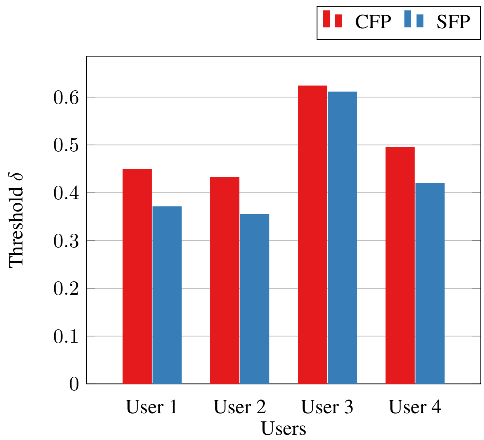
Figure 5.4: Amount of data collected with different policies

Fig. 5.3 compares the energy consumption spent by each user while adopting the CFP and SFP policies. The results are obtained at the end of the evaluation period of 600 timeslots. As expected, User 3 consumes significantly less energy than the other users. Indeed, being at the beginning of the evaluation period her remaining charge of the battery only 10%, the framework tries to limit its contribution. Surprisingly, both CFP and SFP provide a nearly equal energy consumption. Thus, under the current setting, the framework makes the users with low remaining charge to contribute data only when it is sure it brings utility to the collector. The energy consumption of the other users exhibits a similar behavior, being the difference between Users 1 and 2 and User 4 as low as 0.6 J. It should be noted that for a smartphone equipped with a 2550 mAh and a 3.85 V battery such as the Samsung Galaxy s6, the total energy at disposal is as of 35343 J. As a consequence, it is possible to conclude that contributing data opportunistically to a MCS system does not affect dramatically the performance of today mobile devices. For example, the energy consumption a 10 minutes period is as low as 12 J.

Having evaluated the amount of data contributed and the cost to generate such information, now the study is about the impact of the parameters δ and γ on the overall performance. Fig. 5.5 analyzes the behavior of the threshold δ that enables the device to decide whether to perform sensing and reporting. For the analysis, the same setting used to generate the previous results is kept. Fig. 5.5(a) shows the values δ assumes during the evaluation period for all the users, having set the CFP policy. For all the users, δ increases with time. While contributing data, from one hand the users spend energy, having an impact on the remaining charge of the battery (parameter δ_b). On the other hand, delivering data increases the amount of already reported data (parameter δ_d). As expected, for User 3 the threshold δ is always higher than for the other users because her remaining battery charge is the lowest. It is interesting to notice that the curves for Users 1 and 2 are close one each other. The reason is twofold. First, the two users have a similar amount of remaining battery charge at the beginning of the evaluation period (see Table 5.1(a)). Moreover, they contribute nearly the same amount of data, see Fig. 5.1. Fig. 5.5(b) compares the values δ assumes at the end of the evaluation period for both policies CFP and SFP. Similarly to the results obtained for energy consumption (see Fig. 5.3), for all the users but User 3 the difference between the values δ assumes with the two policies remain constant. For User 3, the difference is small. Intuitively, higher values of δ correspond in having low energy consumption.



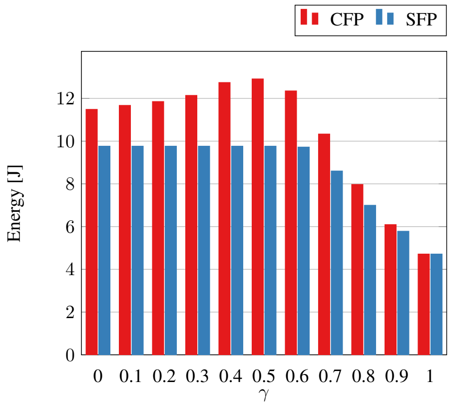
(a) Threshold δ during the period of analysis



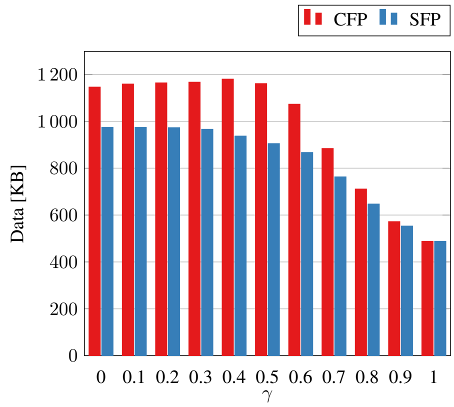
(b) Comparison between data collection policies

Figure 5.5: Analysis of threshold δ

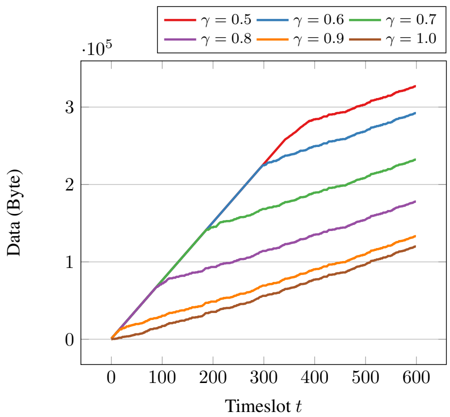
Fig. 5.6 studies the impact of the parameter γ on the performance of the framework. Recalling (4.1) and (4.2), the parameter γ defines the weight of the data collection utility and the smartphone sensing potential for taking decisions to perform sensing and reporting. Fig. 5.6(a) and Fig. 5.6(b) shows respectively the amount of data generated and energy the device consume for both CFP and SFP policies with increasing values of γ. Low values of γ give more importance to the data collection utility, hence the devices contribute data even if they experience a high energy cost as Fig. 5.6(a) highlights. On the other side, with the increase of the parameter γ, the framework is tailored to pay more attention to the potential the device have in perform sensing. This translates in a better energy management at the cost of delivering a lower amount of data. For the CFP policy it is also possible to notice that the energy cost is smoothly increasing for values of γ below 0.5 (see Fig. 5.6(a)) although the amount of data generated remains almost constant (see Fig. 5.6(b)). This energy consumption should be attributed to the utilization context U_s. When the parameter γ assumes values higher than 0.6 both the energy and the amount of collected data have a sharp decrease. As the relevance of the data collection utility decreases, the smartphone sensing potential factor becomes crucial in order to take sensing and reporting decisions. To better analyze such decrease, the plot in Fig. 5.6(c) shows the amount of data User 4 generates for values of $\gamma > 0.6$. Similarly to Fig. 5.1, in all the curves the increase is at first sharp and then becomes smooth. The higher the values γ assumes, the sooner the change of slope and the lower the amount of data collected, which in turns leads to energy savings.



(a) Energy Cost



(b) Comparison amount of data collected



(c) Amount of data collected by User 4

Figure 5.6: Analysis of parameter γ

Chapter 6

CrowdSenSim

CrowdSenSim simulates the process of data collection in a large scale scenario with thousands of participants contributing data in a real city environment. This custom simulator is a discrete-event simulator where the participants of a MCS system contribute data to the collector opportunistically. The objectives of CrowdSenSim consist in measuring the costs the devices experience and determining the amount of contributed data. Furthermore, the simulator is general purpose, so any data collection framework can exploit it for its purposes. Having evaluated the performance of the framework proposed in this thesis analytically, the simulator conducts experiments to investigate the effectiveness of the framework in a large scale scenario. The objective is to assess the efficiency of the algorithm taking into account a large number of participants contributing data in a real city. The following sections present the detailed functionality along with the simulation scenario and results.

6.1 Description and Set-up of the Simulator

The key idea of the simulator is to consider a real city environment and to collect data from thousands of users that walk all around the city. The salient features are as follows:

- The participants move in a realistic city environment.
- The simulator supports pedestrian mobility.
- The data generation process uses latest generation sensors commonly available in mobile devices.

- Mobile devices report the collected data connecting to LTE base station and WiFi hotspots. The real coordinates of the antennas (latitude, longitude and altitude) are used to calculate the distance from users.

Each mobile device generates data opportunistically from a set of sensors including accelerometer, temperature and pressure sensor and transmits data over WiFi. The sensing equipment is the real equipment implemented in current smartphones and tablets. Specifically, this study considers the FXOS8700CQ 3-axis linear accelerometer from Freescale Semiconductor and the BMP280 from Bosch, which is a digital pressure and temperature sensor. Equation (4.9) describes WiFi power consumption spent by the devices for communication. Table 5.1 presents the detailed information of the simulation parameters.

The center of Luxembourg City was selected for the simulations. It covers an area of 1.11 km^2 and is the home of many national and international institutional buildings.

Pedestrian Mobility

By means of DigiPoint [162] from Zonum Solutions 1, it is possible to derive information about the streets of the city in the form of a set of coordinates C containing <*latitude, longitude, altitude*>. The set of coordinates C taken into account is shown in the map of Luxembourg city center in Fig. 6.1. The participants move along the streets of the city, being their original location a randomly assigned coordinate in the set C. The number of participants ranges from 100 to a maximum of 20 000, being the population of Luxembourg city of 107 340 inhabitants as of end 2014. For the sake of simplicity, the start time of the walk is uniformly distributed between 8:00 am and 13:30. Each participant holds one mobile device only and walks for a period of time that is uniformly distributed between [10,30] minutes with an average speed uniformly distributed between [1,1.5] m/s. Once the walk time period expires, the participant stops contributing to the system. This hypothesis provides a worst case estimation, having the users contributing for a little period of time only along the day.

Antennas

The antennas are mapped taking into account both LTE towers and WiFi hotspots in Luxembourg City Center. The position of LTE towers shown in Fig. 6.2 is taken from OpenSignal [163], which is a source of insight into the coverage and performance of Mobile Operators worldwide. The data is crowdsourced by users of the OpenSignal app, downloaded over 15 million times, which constantly monitors the coverage and performance of their mobile connection.

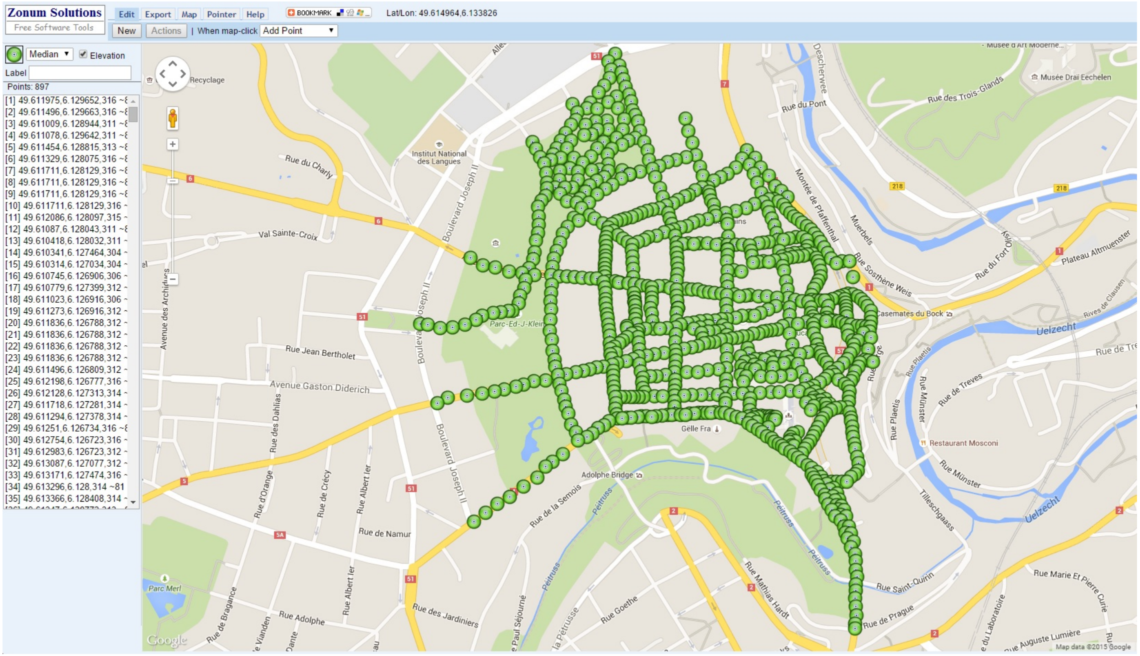


Figure 6.1: Map from DigiPoint

The position of WiFi hotspots in Luxembourg city center is taken from WiFi coverage of HotCity [164] and shown in Fig. 6.3.

6.2 Simulation Results

For evaluation purposes experiments are conducted aiming at assess the amount of information gathered from the devices and the energy consumption they experience while contributing data.

Energy Consumption for Sensing and Reporting

To analyze the consumption the device experience the number of participants is 10 000. For a worst case analysis, the devices generate data continuously during users' movements.

Fig. 6.4(a) analyzes the user distribution for the energy spent for sensing. On average, the users spend 374.617 uAh and in the worst case few users experience a consumption that is nearly double than the average. If compared to the capacity available in today smartphones (in the order of 2000 mAh), it is possible to conclude that the energy consumption for sensing is negligible with respect to the energy spent for communications.

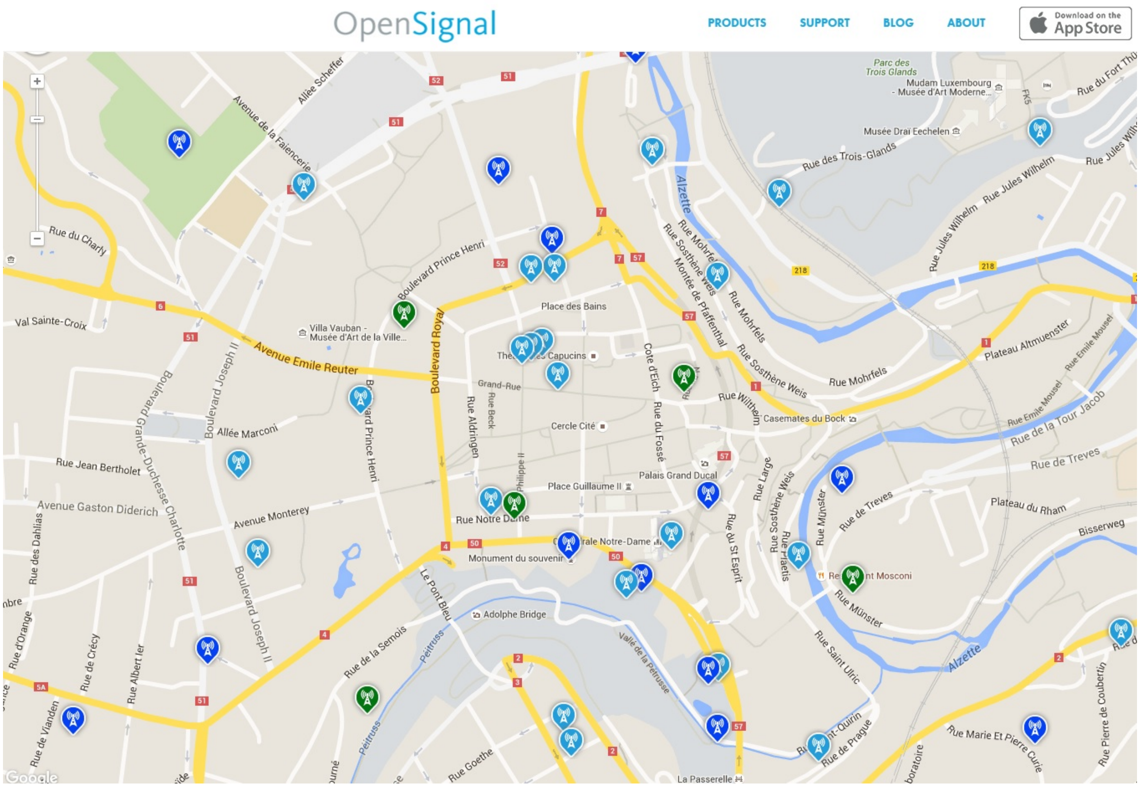


Figure 6.2: Map of LTE base stations from OpenSignal

Fig. 6.4(b) illustrates the user distribution for the energy spent for communications. As only WiFi was considered, the shape of the distribution is similar to the one presented in Fig. 6.4(a).

Amount of Data Collected

To assess the amount of information the device contribute, two types of experiments are conducted. First, the evaluation considers the amount of data each sensor generates. Then, the Luxembourg City center is divided into five areas and in each area it is measured the Sample Distribution (SD). SD is a metric that aims at assessing the distribution of the samples per area. The objective is to understand the amount and the variability of the data generation process along the time period 8:00-14:00. For both experiments, the number of participants was 20 000.

Fig. 6.5 shows the amount of data collected with the increase of the number of participants. The amount of data is proportional to the sampling frequencies of the three sensors. Recalling that each user only contributes for a little period of time (10 to 30 minutes), the amount of collected information is remarkable. For

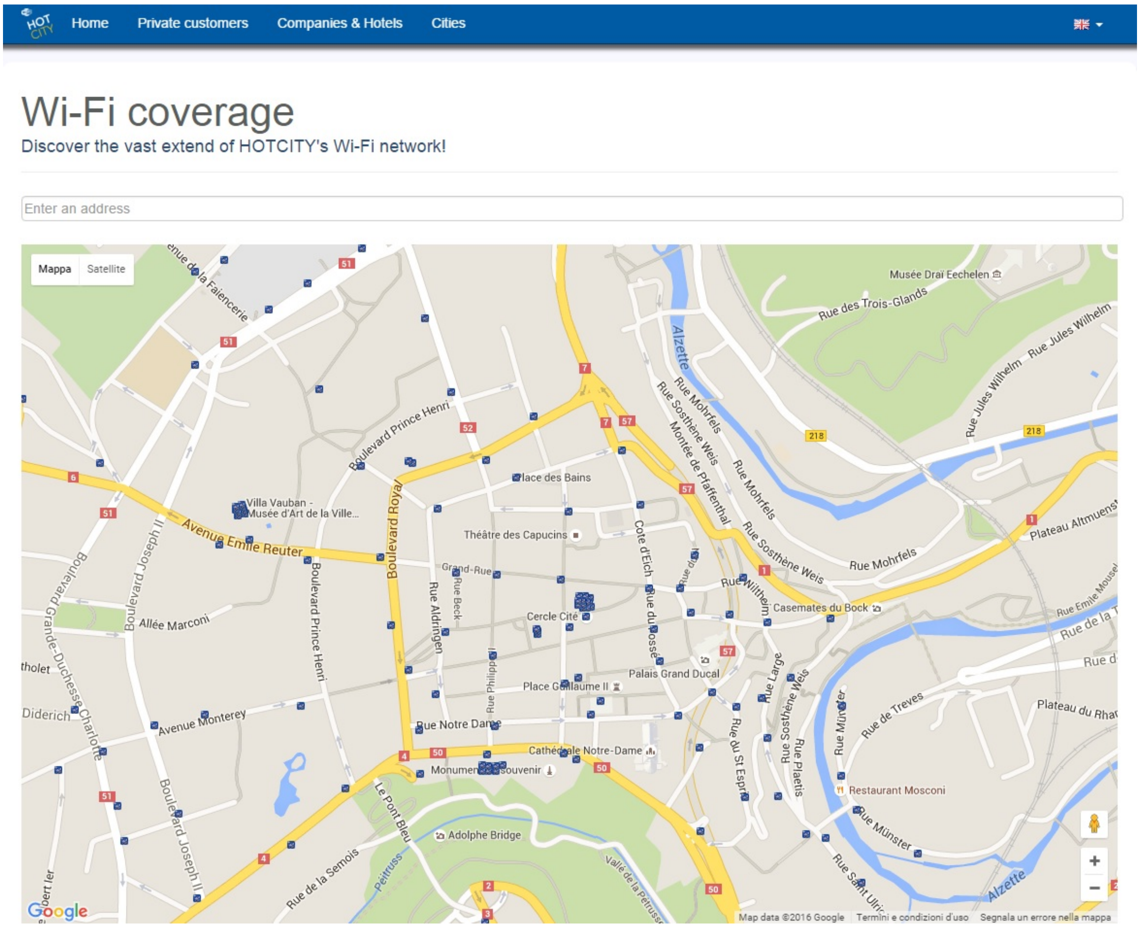


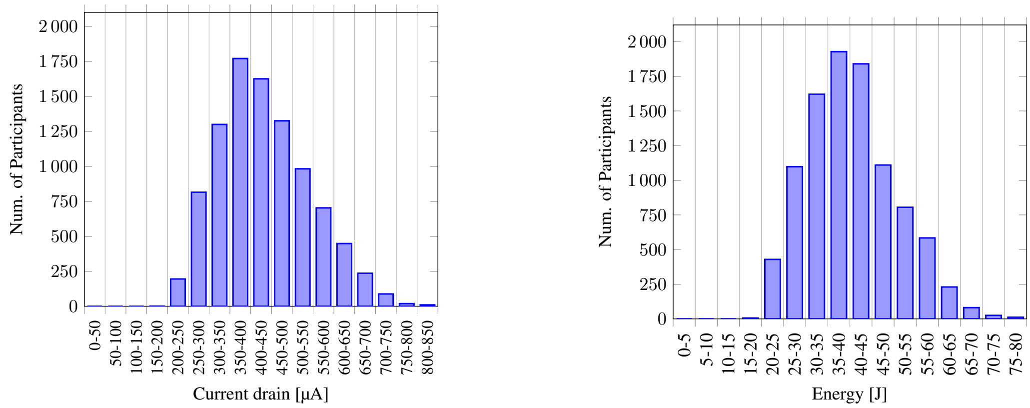
Figure 6.3: Map of WiFi hotspots from HotCity

example, having 5 000 users would generate 783 MB, 3.71 GB and 3.20 GB for the accelerometer, temperature and pressure sensors respectively.

Having the knowledge on the amount of data the users can contribute is important for the applications and to determine the accuracy in mapping a phenomena. However, to draw more precise conclusions it becomes fundamental to determine also where and when the samples are generated. For this reason, it is introduced a new metric called Sample Distribution (SD) to assess the distribution of the sample generation. SD is measured in sample per meter and is defined as follows:

$$\mathrm{SD} = \frac{N_s^a|_t}{\overline{\Delta}}, \tag{6.1}$$

where $\overline{\Delta}$ is the average distance between samples and $N_s^a|_t$ is the number of samples



(a) Sensing Cost

(b) Communication Cost

Figure 6.4: Energy spent for sensing and communication

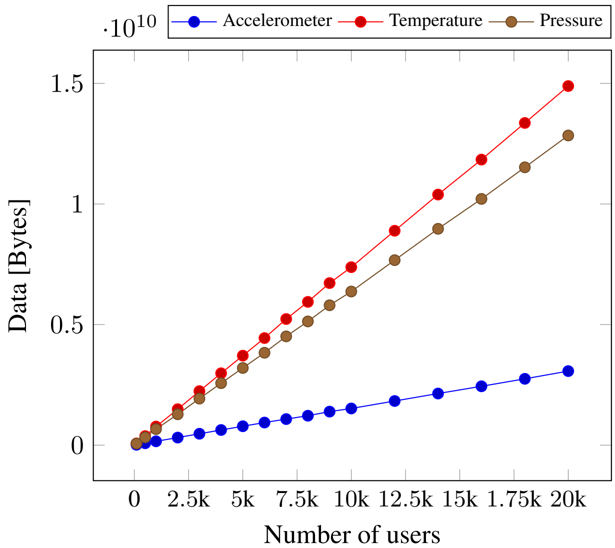


Figure 6.5: Amount of data generated

generated (see Table 4.1). The parameter $\overline{\Delta}$ is defined as follows:

$$\overline{\Delta} = \frac{\sum_{\substack{i,j \\ i \geq j}}^{n} d(i,j)}{\frac{n(n-1)}{2}}. \tag{6.2}$$

The term $d(i,j)$ is the distance (in meters) between the location where the samples i and j were generated.

Fig. 6.6 shows graphically the SD in the entire city center for the time period 10:00-10:30. The considers only accelerometer samples. The SD metric weakly depends on the size of the area. Although being wider than Area 3, a large part of Area 2 is a public park with a fewer number of streets at user disposal. The reason is that the SD metric measures the distribution of the samples taking into account the location where they have been generated, namely a specific coordinate identified by the triple <*latitude, longitude, altitude*>. As a consequence, highly dense areas such as Area 5 exhibits high values of SD.

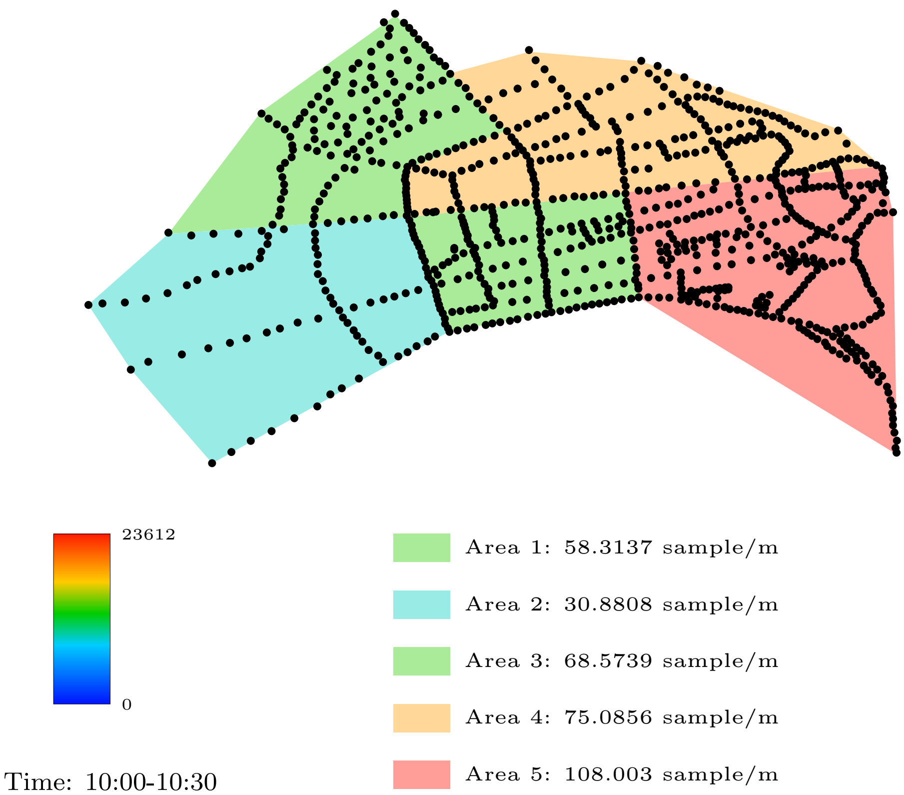


Figure 6.6: Sample distribution (SD) for 20 000 users moving in the center of Luxembourg

Fig. 6.7 shows the distribution of SD for all the 5 areas for the entire simulation period. It is interesting to notice that the lowest values of SD occur for the initial end final time intervals (8:00-8:30 and 13:30-14:00). Being the user uniformly distributed between 8:00 am and 13:30 and moving for at maximum 30 minutes, during the initial and final time interval the number of participants is low.

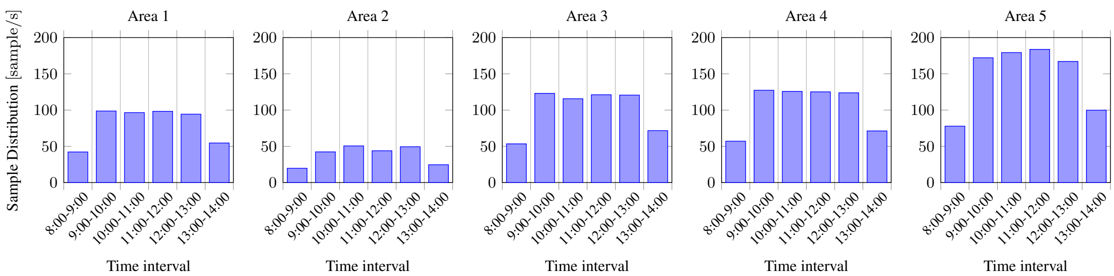


Figure 6.7: Sample distribution (SD) for the different areas over the time period 8:00-14:00

Chapter 7

Conclusion and Future Work

This thesis proposes an optimized distributed framework for data collection in opportunistic MCS systems. The framework aims at minimizing the energy consumption for the participants in performing sensing and reporting while maximizing the utility in data collection for the cloud collector. The performance of the framework is evaluated both analytically and through simulations in a real urban environment with a large number of participants. The thesis analyzes the costs the participants experience and the amount of data the system allows to gather. The analytical results highlight that the major contribution to energy consumption is attributed to reporting and not sensing. The simulation results confirm the effectiveness of the proposed approach in a real urban environment for a large number of participants.

To extend current functionalities of the framework, as future work it is possible to consider caching policies to buffer samples and queuing policies with different schedulers to improve the efficiency of data delivery. In other words, the idea is to include guidelines to implement Quality of Service (QoS) and Age of Information in MCS systems, considering the utility of the cloud collector for packets generated from different sensors and their timestamps. Moreover, it would be interesting to model different environmental contexts (e.g., indoor, outdoor, office, pocket) and the transitions from one to another one. This objective can be achieved exploiting Hidden Markov Models (HMM). For instance, passing from "smarthpone in pocket and user not moving" to "smartphone outside and user moving", all the data collection features explained in the taxonomy change.

Furthermore, MCS is a hot research topic and embraces several interesting research areas. The following list outlines potential future research directions:

- Use of wearable devices to improve smartphone sensing. To analyze these architectures it is fundamental to take into account task assignment in coordinated manner between mobile and wearable devices to not overlap.

Furthemore, energy efficiency is one the most important issues as wearable devices have a very limited amount of resources and batteries drain quickly.

- Connection with other relevant research areas, as *mobile cloud computing*, *big data* and *smart cities*. For instance, there are some existing works connected with cloud computing, but they are more oriented to personal sensing than crowd (e.g., databases in health care with private sensed data that can be stored using mobile devices).
- Collaboration between mobile devices and vehicular networks. It exploits the large availability of embedded sensors in smartphones that cover a huge areas of users that move in cars, buses, etc. Futhermore, it investigates communication techniques which efficiently adapt to the highly volatile and unstable vehicular environment.
- Detection of activities to prevent pedestrian injuries and deaths [165] as road projects, taking into account the needs of all users of the transportation system. It is also strictly connected to the huge research area of public safety, which investigates crimes and natural disasters.
- "Learning" things, like the smart thermometer [166] that automatically adapts as people habits and season changes, with movement sensor and recording activity. It permits energy savings and a great impact in developing smart houses.
- Detection of pollution patterns through machine learning, for gas emission reduction and citizen healthcare. These algorithms exploit known information to recognize specified pollution pattern for environmental monitoring and propose possible solutions [167].

References

[1] *Global mobile statistics 2014 Part A: Mobile subscribers; handset market share; mobile operators*. 2014. URL: http://mobiforge.com/research-analysis/global-mobile-statistics-2014-part-a-mobile-subscribers-handset-market-share-mobile-operators.

[2] *Android on 80 percent of smart phones shipped in 2013*. 2014. URL: http://www.canalys.com/newsroom/android-80-smart-phones-shipped-2013.

[3] *Gartner Says Smartphone Sales Surpassed One Billion Units in 2014*. 2015. URL: http://www.gartner.com/newsroom/id/2996817.

[4] *Global Smartphone Growth Expected to Slow to 11.3 percent in 2015 as Market Penetration Increases in Top Markets, According to IDC*. 2015. URL: http://www.idc.com/getdoc.jsp?containerId=prUS25641615.

[5] Hans Schaffers, Nicos Komninos, Marc Pallot, Brigitte Trousse, Michael Nilsson, and Alvaro Oliveira. "Smart Cities and the Future Internet: Towards Cooperation Frameworks for Open Innovation". English. In: *The Future Internet*. Ed. by John Domingue, Alex Galis, Anastasius Gavras, Theodore Zahariadis, Dave Lambert, Frances Cleary, Petros Daras, Srdjan Krco, Henning Müller, Man-Sze Li, Hans Schaffers, Volkmar Lotz, Federico Alvarez, Burkhard Stiller, Stamatis Karnouskos, Susanna Avessta, and Michael Nilsson. Vol. 6656. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2011, pp. 431–446. ISBN: 978-3-642-20897-3. DOI: 10.1007/978-3-642-20898-0_31. URL: http://dx.doi.org/10.1007/978-3-642-20898-0_31.

[6] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi. "Internet of Things for Smart Cities". In: *IEEE Internet of Things Journal* 1.1 (2014), pp. 22–32. ISSN: 2327-4662. DOI: 10.1109/JIOT.2014.2306328.

[7] Jakob Eriksson, Lewis Girod, Bret Hull, Ryan Newton, Samuel Madden, and Hari Balakrishnan. "The pothole patrol: using a mobile sensor network for road surface monitoring". In: *Proceedings of the 6th international conference on Mobile systems, applications, and services*. ACM. 2008, pp. 29–39.

[8] Sarfraz Nawaz, Christos Efstratiou, and Cecilia Mascolo. "ParkSense: A Smartphone Based Sensing System for On-street Parking". In: *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*. MobiCom '13. Miami, Florida, USA: ACM, 2013, pp. 75–86. ISBN: 978-1-4503-1999-7. DOI: `10.1145/2500423.2500438`. URL: `http://doi.acm.org/10.1145/2500423.2500438`.

[9] Suhas Mathur, Tong Jin, Nikhil Kasturirangan, Janani Chandrasekaran, Wenzhi Xue, Marco Gruteser, and Wade Trappe. "Parknet: drive-by sensing of road-side parking statistics". In: *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM. 2010, pp. 123–136.

[10] Xiang Chen, Yiran Chen, Mian Dong, and C. Zhang. "Demystifying energy usage in smartphones". In: *51st ACM/EDAC/IEEE Design Automation Conference (DAC)*. 2014, pp. 1–5.

[11] 6sensorlabs. *Gluten sensor*. [Online]. 2015.

[12] David Hasenfratz, Olga Saukh, Silvan Sturzenegger, and Lothar Thiele. "Participatory air pollution monitoring using smartphones". In: *Mobile Sensing* (2012).

[13] Matthias Budde, Pierre Barbera, Rayan El Masri, Till Riedel, and Michael Beigl. "Retrofitting Smartphones to Be Used As Particulate Matter Dosimeters". In: *Proceedings of the 2013 International Symposium on Wearable Computers*. ISWC '13. Zurich, Switzerland: ACM, 2013, pp. 139–140. ISBN: 978-1-4503-2127-3. DOI: `10.1145/2493988.2494342`. URL: `http://doi.acm.org/10.1145/2493988.2494342`.

[14] Yasunori Ishigaki, Yuki Matsumoto, Ryo Ichimiya, and Kiyoshi Tanaka. "Development of mobile radiation monitoring system utilizing smartphone and its field tests in Fukushima". In: *IEEE Sensors Journal* 13.10 (2013), pp. 3520–3526.

[15] Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L Littman. "Activity recognition from accelerometer data". In: *AAAI*. Vol. 5. 2005, pp. 1541–1546.

[16] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. "Activity recognition using cell phone accelerometers". In: *ACM SigKDD Explorations Newsletter* 12.2 (2011), pp. 74–82.

[17] Derick Johnson, Mohan M Trivedi, et al. "Driving style recognition using a smartphone as a sensor platform". In: *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on*. IEEE. 2011, pp. 1609–1615.

[18] Christian Schuss, Tore Leikanger, Bernd Eichberger, and Timo Rahkonen. "Efficient use of solar chargers with the help of ambient light sensors on smartphones". In: *16th Conference of Open Innovations Association (FRUCT16)*. IEEE. 2014, pp. 79–85.

[19] Yi F Dong, S Kanhere, Chun Tung Chou, and Nirupama Bulusu. "Automatic collection of fuel prices from a network of mobile cameras". In: *Distributed computing in sensor systems*. Springer, 2008, pp. 140–156.

[20] Sasank Reddy, Andrew Parker, Josh Hyman, Jeff Burke, Deborah Estrin, and Mark Hansen. "Image browsing, processing, and clustering for participatory sensing: lessons from a DietSense prototype". In: *Proceedings of the 4th workshop on Embedded networked sensors*. ACM. 2007, pp. 13–17.

[21] M. Kok, J.D. Hol, T.B. Schon, F. Gustafsson, and H. Luinge. "Calibration of a magnetometer in combination with inertial sensors". In: *15th International Conference on Information Fusion (FUSION)*. 2012, pp. 787–793.

[22] Anna-Kaisa Pietiläinen, Earl Oliver, Jason LeBrun, George Varghese, and Christophe Diot. "MobiClique: middleware for mobile social networking". In: *Proceedings of the 2nd ACM workshop on Online social networks*. ACM. 2009, pp. 49–54.

[23] Lei Zhang, Jiangchuan Liu, Hongbo Jiang, and Yong Guan. "SensTrack: Energy-Efficient Location Tracking With Smartphone Sensors". In: *IEEE Sensors Journal* 13.10 (2013), pp. 3775–3784. ISSN: 1530-437X. DOI: 10.1109/JSEN.2013.2274074.

[24] Immanuel Schweizer, Roman Bärtl, Axel Schulz, Florian Probst, and Max Mühläuser. "NoiseMap-real-time participatory noise maps". In: *Proc. 2nd Int'l Workshop on Sensing Applications on Mobile Phones (PhoneSense'11)*. 2011, pp. 1–5.

[25] Armir Bujari, Bogdan Licar, and Claudio E Palazzi. "Movement pattern recognition through smartphone's accelerometer". In: *IEEE Consumer communications and networking conference (CCNC)*. IEEE. 2012, pp. 502–506.

[26] Ilari Shafer and Mark L Chang. "Movement detection for power-efficient smartphone WLAN localization". In: *Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems*. ACM. 2010, pp. 81–90.

[27] Pekka Siirtola and Juha Röning. "Recognizing human activities user-independently on smartphones based on accelerometer data". In: *International Journal of Interactive Multimedia and Artificial Intelligence* 1.5 (2012).

[28] Ashiq Anjum and Muhammad Usman Ilyas. "Activity recognition using smartphone sensors". In: *IEEE Consumer Communications and Networking Conference (CCNC)*. IEEE. 2013, pp. 914–919.

[29] Mohammed Shoaib, Hans Scholten, and Paul JM Havinga. "Towards physical activity recognition using smartphone sensors". In: *IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC) Ubiquitous Intelligence and Computing*. IEEE. 2013, pp. 80–87.

[30] Jiangpeng Dai, Xiaole Bai, Zhimin Yang, Zhaohui Shen, and Dong Xuan. "PerFallD: A pervasive fall detection system using mobile phones". In: *8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*. IEEE. 2010, pp. 292–297.

[31] Yi He, Ye Li, and Shu-Oi Bao. "Fall Detection by built-in tri-accelerometer of smartphone". In: *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*. IEEE. 2012, pp. 184–187.

[32] Mohammad Ashfak Habib, Mas S Mohktar, Shahrul Bahyah Kamaruzzaman, Kheng Seang Lim, Tan Maw Pin, and Fatimah Ibrahim. "Smartphone-based solutions for fall detection and prevention: challenges and open issues". In: *Sensors* 14.4 (2014), pp. 7181–7208.

[33] Shuangquan Wang, Canfeng Chen, and Jian Ma. "Accelerometer based transportation mode recognition on mobile phones". In: *Asia-Pacific Conference on Wearable Computing Systems (APWCS)*. IEEE. 2010, pp. 44–46.

[34] Sasank Reddy, Jeff Burke, Deborah Estrin, Mark Hansen, and Mani Srivastava. "Determining transportation mode on mobile phones". In: *12th IEEE International Symposium on Wearable Computers (ISWC)*. IEEE. 2008, pp. 25–28.

[35] Samuli Hemminki, Petteri Nurmi, and Sasu Tarkoma. "Accelerometer-based transportation mode detection on smartphones". In: *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. ACM. 2013, p. 13.

[36] Sasank Reddy, Min Mun, Jeff Burke, Deborah Estrin, Mark Hansen, and Mani Srivastava. "Using mobile phones to determine transportation modes". In: *ACM Transactions on Sensor Networks (TOSN)* 6.2 (2010), p. 13.

[37] Dongyoun Shin, Daniel Aliaga, Bige Tunçer, Stefan Müller Arisona, Sungah Kim, Dani Zünd, and Gerhard Schmitt. "Urban sensing: Using smartphones for transportation mode classification". In: *Computers, Environment and Urban Systems* (2014).

[38] Hao Xia, Yanyou Qiao, Jun Jian, and Yuanfei Chang. "Using smart phone sensors to detect transportation modes". In: *Sensors* 14.11 (2014), pp. 20843–20865.

[39] Muhammad Shoaib, Stephan Bosch, Ozlem Durmaz Incel, Hans Scholten, and Paul JM Havinga. "Fusion of smartphone motion sensors for physical activity recognition". In: *Sensors* 14.6 (2014), pp. 10146–10176.

[40] Oliver Walter, Joerg Schmalenstroeer, Andreas Engler, and Reinhold Haeb-Umbach. "Smartphone-based sensor fusion for improved vehicular navigation". In: *Positioning Navigation and Communication (WPNC), 2013 10th Workshop on*. IEEE. 2013, pp. 1–6.

[41] Yan Michalevsky, Dan Boneh, and Gabi Nakibly. "Gyrophone: Recognizing speech from gyroscope signals". In: *Proc. 23rd USENIX Security Symposium (SEC'14), USENIX Association*. 2014.

[42] Wonho Kang, Seongho Nam, Youngnam Han, and Sookjin Lee. "Improved heading estimation for smartphone-based indoor positioning systems". In: *IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. IEEE. 2012, pp. 2449–2453.

[43] F Keller, Thomas Willemsen, and Harald Sternberg. "Calibration of smartphones for the use in indoor navigation". In: *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE. 2012, pp. 1–8.

[44] *WeatherSignal: Use your mobile as a weather station*. 2013. URL: http://weathersignal.com/.

[45] Aart Overeem, JC R Robinson, Hidde Leijnse, Gert-Jan Steeneveld, BK P Horn, and Remko Uijlenhoet. "Crowdsourcing urban air temperatures from smartphone battery temperatures". In: *Geophysical Research Letters* 40.15 (2013), pp. 4081–4085. DOI: 10.1002/grl.50786.

[46] Guangwen Liu, Khan Muhammad Asif Hossain, Masayuki Iwai, Masaki Ito, Yoshito Tobe, Kaoru Sezaki, and Dunstan Matekenya. "Beyond horizontal location context: measuring elevation using smartphone's barometer". In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM. 2014, pp. 459–468.

[47] Kartik Sankaran, Minhui Zhu, Xiang Fa Guo, Akkihebbal L Ananda, Mun Choon Chan, and Li-Shiuan Peh. "Using mobile phone barometer for low-power transportation context detection". In: *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*. ACM. 2014, pp. 191–205.

[48] Haibo Ye, Tao Gu, Xianping Tao, and Jian Lu. "B-Loc: scalable floor localization using barometer on smartphone". In: *IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. IEEE. 2014, pp. 127–135.

[49] Salvatore Vanini and Stefano Giordano. "Adaptive context-agnostic floor transition detection on smart mobile devices". In: *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*. IEEE. 2013, pp. 2–7.

[50] Keisuke Komeda, Masahiro Mochizuki, and Nobuhiko Nishiko. "User activity recognition method based on atmospheric pressure sensing". In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM. 2014, pp. 737–746.

[51] Kartik Muralidharan, Azeem Javed Khan, Archan Misra, Rajesh Krishna Balan, and Sharad Agarwal. "Barometric phone sensors: more hype than hope!" In: *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications*. ACM. 2014, p. 12.

[52] Ammar El Halabi and Hassan Artail. "Integrating pressure and accelerometer sensing for improved activity recognition on smartphones". In: *Third International Conference on Communications and Information Technology (ICCIT)*. IEEE. 2013, pp. 121–125.

[53] *How phone batteries measure the weather*. 2013. URL: `http://opensignal.com/reports/battery-temperature-weather/`.

[54] Siamak Aram, Amedeo Troiano, and Eros Pasero. "Environment sensing using smartphone". In: *IEEE Sensors Applications Symposium (SAS)*. IEEE. 2012, pp. 1–4.

[55] J. Weppner and P. Lukowicz. "Bluetooth based collaborative crowd density estimation with mobile phones". In: *2013 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 2013, pp. 193–200. DOI: `10.1109/PerCom.2013.6526732`.

[56] Xiang Sheng, Jian Tang, and Weiyi Zhang. "Energy-efficient collaborative sensing with mobile phones". In: *Proceedings IEEE INFOCOM*. 2012, pp. 1916–1924. DOI: `10.1109/INFCOM.2012.6195568`.

[57] Beom-Ju Shin, Kwang-Won Lee, Sun-Ho Choi, Joo-Yeon Kim, Woo Jin Lee, and Hyung Seok Kim. "Indoor WiFi positioning system for Android-based smartphone". In: *Int'l Conference on Information and Communication Technology Convergence*. 2010, pp. 319–320.

[58] Azade Farshad, Mahesh K Marina, and Francisco Garcia. "Urban wifi characterization via mobile crowdsensing". In: *Network Operations and Management Symposium (NOMS), 2014 IEEE*. IEEE. 2014, pp. 1–9.

[59] Roy Want. "Near field communication". In: *IEEE Pervasive Computing* 3 (2011), pp. 4–7.

[60] Esko Strommer, Jouni Kaartinen, Juha Parkka, Arto Ylisaukko-Oja, and Ilkka Korhonen. "Application of near field communication for health monitoring in daily life". In: *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*. IEEE. 2006, pp. 3246–3249.

[61] T.-T.T. Pham and J.C. Ho. "What are the core drivers in consumer adoption of NFC-based mobile payments?: A proposed research framework". In: *International Conference on Management of Engineering Technology (PICMET)*. 2014, pp. 3041–3049.

[62] John Chon and Hojung Cha. "Lifemap: A smartphone-based context provider for location-based services". In: *IEEE Pervasive Computing* 2 (2011), pp. 58–67.

[63] Thomas Olutoyin Oshin, Stefan Poslad, and Andy Ma. "Improving the Energy-Efficiency of GPS based location sensing smartphone applications". In: *IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. IEEE. 2012, pp. 1698–1705.

[64] Jeongyeup Paek, Joongheon Kim, and Ramesh Govindan. "Energy-efficient rate-adaptive GPS-based positioning for smartphones". In: *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM. 2010, pp. 299–314.

[65] Alex Varshavsky, Mike Y Chen, Eyal De Lara, Jon Froehlich, Daniel Haehnel, Jeffrey Hightower, Antonio Lamarca, Fred Potter, Timothy Sohn, Ke Tang, et al. "Are GSM phones THE solution for localization?" In: *7th IEEE Workshop on Mobile Computing Systems and Applications (WMCSA)*. IEEE. 2005, pp. 34–42.

[66] Muhammad Yasir, Siu-Wai Ho, and Badri N Vellambi. "Indoor positioning system using visible light and accelerometer". In: *Journal of Lightwave Technology* 32.19 (2014), pp. 3306–3316.

[67] Pengfei Zhou, Yuanqing Zheng, Zhenjiang Li, Mo Li, and Guobin Shen. "Iodetector: A generic service for indoor outdoor detection". In: *Proceedings of the 10th acm conference on embedded network sensor systems*. ACM. 2012, pp. 113–126.

[68] *SensorWiki.org*. 2014. URL: `http://sensorwiki.org/doku.php`.

[69] Jae-Yeon Won, Hyunsurk Ryu, Tobi Delbruck, Jun Haeng Lee, and Jiang Hu. "Proximity Sensing Based on a Dynamic Vision Sensor for Mobile Devices". In: *IEEE Transactions on Industrial Electronics* 62.1 (2015), pp. 536–544.

[70] Antti J Eronen, Vesa T Peltonen, Juha T Tuomi, Anssi P Klapuri, Seppo Fagerlund, Timo Sorsa, Gaëtan Lorho, and Jyri Huopaniemi. "Audio-based context recognition". In: *IEEE Transactions on Audio, Speech, and Language Processing* 14.1 (2006), pp. 321–329.

[71] Mattia Rossi, Sebastian Feese, Oliver Amft, Nils Braune, Sandro Martis, and G Troster. "AmbientSense: A real-time ambient sound recognition system for smartphones". In: *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*. IEEE. 2013, pp. 230–235.

[72] Annamaria Mesaros, Toni Heittola, Antti Eronen, and Tuomas Virtanen. "Acoustic event detection in real life recordings". In: *18th European Signal Processing Conference*. Citeseer. 2010, pp. 1267–1271.

[73] Hong Lu, Wei Pan, Nicholas D Lane, Tanzeem Choudhury, and Andrew T Campbell. "SoundSense: scalable sound sensing for people-centric applications on mobile phones". In: *Proceedings of the 7th international conference on Mobile systems, applications, and services*. ACM. 2009, pp. 165–178.

[74] Hong Lu, AJ Bernheim Brush, Bodhi Priyantha, Amy K Karlson, and Jie Liu. "SpeakerSense: energy efficient unobtrusive speaker identification on mobile phones". In: *Pervasive Computing*. Springer, 2011, pp. 188–205.

[75] Chunming Gao, Fanyu Kong, and Jindong Tan. "Healthaware: Tackling obesity with health aware smart phone systems". In: *Robotics and Biomimetics (ROBIO), 2009 IEEE International Conference on*. Ieee. 2009, pp. 1549–1554.

[76] Prashanth Mohan, Venkata N Padmanabhan, and Ramachandran Ramjee. "Nericell: rich monitoring of road and traffic conditions using mobile smartphones". In: *Proceedings of the 6th ACM conference on Embedded network sensor systems*. ACM. 2008, pp. 323–336.

[77] Emmanouil Koukoumidis, Margaret Martonosi, and Li-Shiuan Peh. "Leveraging smartphone cameras for collaborative road advisories". In: *IEEE Transactions on Mobile Computing* 11.5 (2012), pp. 707–723.

[78] Laura Ruotsalainen, Heidi Kuusniemi, and Ruizhi Chen. "Heading change detection for indoor navigation with a smartphone camera". In: *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE. 2011, pp. 1–7.

[79] Martin Werner, Moritz Kessel, and Chadly Marouane. "Indoor positioning using smartphone camera". In: *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE. 2011, pp. 1–6.

[80] Victor Fragoso, Steffen Gauglitz, Shane Zamora, Jim Kleban, and Matthew Turk. "Translatar: A mobile augmented reality translator". In: *IEEE Workshop on Applications of Computer Vision (WACV)*. IEEE. 2011, pp. 497–502.

[81] Prabal Dutta, Paul M Aoki, Neil Kumar, Alan Mainwaring, Chris Myers, Wesley Willett, and Allison Woodruff. "Common sense: participatory urban sensing using a network of handheld air quality monitors". In: *Proceedings of the 7th ACM conference on embedded networked sensor systems*. ACM. 2009, pp. 349–350.

[82] N.D. Lane, E. Miluzzo, Hong Lu, D. Peebles, T. Choudhury, and A.T. Campbell. "A survey of mobile phone sensing". In: *IEEE Communications Magazine* 48.9 (2010), pp. 140–150. ISSN: 0163-6804. DOI: 10.1109/MCOM.2010.5560598.

[83] W.Z. Khan, Yang Xiang, M.Y. Aalsalem, and Q. Arshad. "Mobile Phone Sensing Systems: A Survey". In: *IEEE Communications Surveys Tutorials* 15.1 (2013), pp. 402–427. ISSN: 1553-877X. DOI: 10.1109/SURV.2012.031412.00077.

[84] X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, and X. Mao. "Incentives for Mobile Crowd Sensing: A Survey". In: *IEEE Communications Surveys Tutorials* PP.99 (2015), pp. 1–1. ISSN: 1553-877X. DOI: 10.1109/COMST.2015.2415528.

[85] L.G. Jaimes, I.J. Vergara-Laurens, and A. Raij. "A Survey of Incentive Techniques for Mobile Crowd Sensing". In: *IEEE Internet of Things Journal* 2.5 (2015), pp. 370–380. ISSN: 2327-4662. DOI: 10.1109/JIOT.2015.2409151.

[86] Delphine Christin, Andreas Reinhardt, Salil S Kanhere, and Matthias Hollick. "A survey on privacy in mobile participatory sensing applications". In: *Journal of Systems and Software* 84.11 (2011), pp. 1928–1946.

[87] Layla Pournajaf, Li Xiong, Daniel A Garcia-Ulloa, and Vaidy Sunderam. *A survey on privacy in mobile crowd sensing task management*. Tech. rep. Technical Report TR-2014-002, Department of Mathematics and Computer Science, Emory University, 2014.

[88] Junxian Huang, Feng Qian, Alexandre Gerber, Z. Morley Mao, Subhabrata Sen, and Oliver Spatscheck. "A Close Examination of Performance and Power Characteristics of 4G LTE Networks". In: *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*. MobiSys '12. Low Wood Bay, Lake District, UK: ACM, 2012, pp. 225–238. ISBN: 978-1-4503-1301-8. DOI: 10.1145/2307636.2307658. URL: http://doi.acm.org/10.1145/2307636.2307658.

[89] R. Friedman, A Kogan, and Y. Krivolapov. "On Power and Throughput Tradeoffs of WiFi and Bluetooth in Smartphones". In: *IEEE Transactions on Mobile Computing* 12.7 (2013), pp. 1363–1376. ISSN: 1536-1233. DOI: 10.1109/TMC.2012.117.

[90] R.K. Ganti, Fan Ye, and Hui Lei. "Mobile crowdsensing: current state and future challenges". In: *IEEE Communications Magazine* 49.11 (2011), pp. 32–39. ISSN: 0163-6804. DOI: 10.1109/MCOM.2011.6069707.

[91] Bin Guo, Zhiwen Yu, Daqing Zhang, and Xingshe Zhou. "From Participatory Sensing to Mobile Crowd Sensing". In: *CoRR* abs/1401.3090 (2014). URL: http://arxiv.org/abs/1401.3090.

[92] Hossein Falaki, Dimitrios Lymberopoulos, Ratul Mahajan, Srikanth Kandula, and Deborah Estrin. "A First Look at Traffic on Smartphones". In: *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*. IMC '10. Melbourne, Australia: ACM, 2010, pp. 281–287. ISBN: 978-1-4503-0483-2. DOI: 10.1145/1879141.1879176. URL: http://doi.acm.org/10.1145/1879141.1879176.

[93] Hoang T. Dinh, Chonho Lee, Dusit Niyato, and Ping Wang. "A survey of mobile cloud computing: architecture, applications, and approaches". In: *Wireless Communications and Mobile Computing* 13.18 (2013), pp. 1587–1611. ISSN: 1530-8677. DOI: 10.1002/wcm.1203. URL: http://dx.doi.org/10.1002/wcm.1203.

[94] *Worldwide sales of smartphones*. Gartner. 2015. URL: http://www.gartner.com/newsroom/id/2996817.

[95] *Worldwide sales of wearables*. Gartner. 2015. URL: http://www.gartner.com/newsroom/id/2913318.

[96] BBC Research. *Wearable Computing: Technologies, Applications and Global Markets*. Feb. 2014.

[97] Jiajun Sun. "Incentive Mechanisms for Mobile Crowd Sensing: Current States and Challenges of Work". In: *CoRR* abs/1310.8364 (2013). URL: http://arxiv.org/abs/1310.8364.

[98] *Smart Cities Will Use 1.1 Billion Connected Things in 2015*. Gartner. 2015. URL: http://www.gartner.com/newsroom/id/3008917.

[99] K. Farkas, G. Feher, A. Benczur, and C. Sidlo. "Crowdsending based public transport information service in smart cities". In: *IEEE Communications Magazine* 53.8 (2015), pp. 158–165. ISSN: 0163-6804. DOI: 10.1109/MCOM.2015.7180523.

[100] M.V. Moreno, A.F. Skarmeta, and A.J. Jara. "How to intelligently make sense of real data of smart cities". In: *International Conference on Recent Advances in Internet of Things (RIoT)*. 2015, pp. 1–6. DOI: 10.1109/RIOT.2015.7104899.

[101] Nicolas Maisonneuve, Matthias Stevens, Maria E Niessen, and Luc Steels. "NoiseTube: Measuring and mapping noise pollution with mobile phones". In: *Information Technologies in Environmental Engineering*. Springer, 2009, pp. 215–228.

[102] Bin Guo, Zhu Wang, Zhiwen Yu, Yu Wang, Neil Y. Yen, Runhe Huang, and Xingshe Zhou. "Mobile Crowd Sensing and Computing: The Review of an Emerging Human-Powered Sensing Paradigm". In: *ACM Comput. Surv.* 48.1 (Aug. 2015), 7:1–7:31. ISSN: 0360-0300. DOI: 10.1145/2794400. URL: http://doi.acm.org.proxy.bnl.lu/10.1145/2794400.

[103] Yang Han, Yanmin Zhu, and Jiadi Yu. "A Distributed Utility-Maximizing Algorithm for Data Collection in Mobile Crowd Sensing". In: *Proceedings of IEEE GLOBECOM*. Austin, USA, Dec. 2014, pp. 277–282.

[104] Huadong Ma, Dong Zhao, and Peiyan Yuan. "Opportunities in mobile crowd sensing". In: *IEEE Communications Magazine* 52.8 (2014), pp. 29–35. ISSN: 0163-6804. DOI: 10.1109/MCOM.2014.6871666.

[105] P.P. Jayaraman, C. Perera, D. Georgakopoulos, and A. Zaslavsky. "Efficient opportunistic sensing using mobile collaborative platform MOSDEN". In: *9th International Conference Conference on Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom)*. 2013, pp. 77–86.

[106] Min Mun, Sasank Reddy, Katie Shilton, Nathan Yau, Jeff Burke, Deborah Estrin, Mark Hansen, Eric Howard, Ruth West, and Péter Boda. "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*. ACM. 2009, pp. 55–68.

[107] Rajib Kumar Rana, Chun Tung Chou, Salil S Kanhere, Nirupama Bulusu, and Wen Hu. "Ear-phone: an end-to-end participatory urban noise mapping system". In: *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*. ACM. 2010, pp. 105–116.

[108] Pengfei Zhou, Yuanqing Zheng, and Mo Li. "How long to wait?: predicting bus arrival time with mobile phone based participatory sensing". In: *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM. 2012, pp. 379–392.

[109] Nirupama Bulusu, Chun Tung Chou, Salil Kanhere, Yifei Dong, Shitiz Sehgal, David Sullivan, and Lupco Blazeski. “Participatory sensing in commerce: Using mobile camera phones to track market price dispersion”. In: *Proceedings of the International Workshop on Urban, Community, and Social Applications of Networked Sensing Systems (UrbanSense 2008)*. 2008, pp. 6–10.

[110] Joaquin Ballesteros, Bogdan Carbunar, Mosaddequr Rahman, Naphtali Rishe, and SS Iyengar. “Towards safe cities: A mobile and social networking approach”. In: *Parallel and Distributed Systems, IEEE Transactions on* 25.9 (2014), pp. 2451–2462.

[111] John Hicks, Nithya Ramanathan, Donnie Kim, Mohamad Monibi, Joshua Selsky, Mark Hansen, and Deborah Estrin. “AndWellness: an open mobile system for activity and experience sampling”. In: *Wireless Health 2010*. ACM. 2010, pp. 34–43.

[112] Emiliano Miluzzo, Cory T Cornelius, Ashwin Ramaswamy, Tanzeem Choudhury, Zhigang Liu, and Andrew T Campbell. “Darwin phones: the evolution of sensing and inference on mobile phones”. In: *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM. 2010, pp. 5–20.

[113] Nikolaos Vastardis and Kun Yang. “Mobile social networks: Architectures, social properties, and key research challenges”. In: *Communications Surveys & Tutorials, IEEE* 15.3 (2013), pp. 1355–1371.

[114] Shitiz Sehgal, Salil S Kanhere, and Chun Tung Chou. “Mobishop: Using mobile phones for sharing consumer pricing information”. In: *Demo Session of the Intl. Conference on Distributed Computing in Sensor Systems*. Citeseer. 2008.

[115] Linda Deng and Landon P Cox. “Livecompare: grocery bargain hunting through participatory sensing”. In: *Proceedings of the 10th workshop on Mobile Computing Systems and Applications*. ACM. 2009, p. 4.

[116] Victor Pankratius, Frank Lind, Anthea Coster, Philip Erickson, and Joshua Semeter. “Mobile crowd sensing in space weather monitoring: the mahali project”. In: *Communications Magazine, IEEE* 52.8 (2014), pp. 22–28.

[117] Kewei Sha, Guoxing Zhan, Weisong Shi, Mark Lumley, Clairy Wiholm, and Bengt Arnetz. “SPA: a smart phone assisted chronic illness self-management system with participatory sensing”. In: *Proceedings of the 2nd International Workshop on Systems and Networking Support for Health Care and Assisted Living Environments*. ACM. 2008, p. 5.

[118] Guanling Chen, Bo Yan, Minho Shin, David Kotz, and Ethan Berke. “MPCS: Mobile-phone based patient compliance system for chronic illness care”. In: *Mobile and Ubiquitous Systems: Networking & Services, MobiQuitous, 2009. MobiQuitous' 09. 6th Annual International*. IEEE. 2009, pp. 1–7.

[119] Jules White, Chris Thompson, Hamilton Turner, Brian Dougherty, and Douglas C Schmidt. “Wreckwatch: Automatic traffic accident detection and notification with smartphones”. In: *Mobile Networks and Applications* 16.3 (2011), pp. 285–303.

[120] Arvind Thiagarajan, Lenin Ravindranath, Katrina LaCurts, Samuel Madden, Hari Balakrishnan, Sivan Toledo, and Jakob Eriksson. “VTrack: accurate, energy-aware road traffic delay estimation using mobile phones”. In: *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*. ACM. 2009, pp. 85–98.

[121] Rita Tse, Yubin Xiao, Giovanni Pau, Marco Roccetti, Serge Fdida, and Gustavo Marfia. “On the Feasibility of Social Network-based Pollution Sensing in ITSs”. In: *arXiv preprint arXiv:1411.6573* (2014).

[122] Emiliano Miluzzo, Nicholas D Lane, Kristóf Fodor, Ronald Peterson, Hong Lu, Mirco Musolesi, Shane B Eisenman, Xiao Zheng, and Andrew T Campbell. “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*. ACM. 2008, pp. 337–350.

[123] Kiran K Rachuri, Mirco Musolesi, Cecilia Mascolo, Peter J Rentfrow, Chris Longworth, and Andrius Aucinas. “EmotionSense: a mobile phones based adaptive platform for experimental social psychology research”. In: *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM. 2010, pp. 281–290.

[124] Shravan Gaonkar, Jack Li, Romit Roy Choudhury, Landon Cox, and Al Schmidt. “Micro-blog: sharing and querying content through mobile phones and social participation”. In: *Proceedings of the 6th international conference on Mobile systems, applications, and services*. ACM. 2008, pp. 174–186.

[125] Kiran K Rachuri, Cecilia Mascolo, Mirco Musolesi, and Peter J Rentfrow. “Sociablesense: exploring the trade-offs of adaptive sampling and computation offloading for social sensing”. In: *Proceedings of the 17th annual international conference on Mobile computing and networking*. ACM. 2011, pp. 73–84.

[126] Xuan Bao and Romit Roy Choudhury. “Movi: mobile phone based video highlights via collaborative sensing”. In: *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM. 2010, pp. 357–370.

[127] Nathan Eagle and Alex Pentland. "Social serendipity: Mobilizing social software". In: *Pervasive Computing, IEEE* 4.2 (2005), pp. 28–34.

[128] Aaron Beach, Mike Gartrell, Sirisha Akkala, Jack Elston, John Kelley, Keisuke Nishimoto, Baishakhi Ray, Sergei Razgulin, Karthik Sundaresan, Bonnie Surendar, et al. "Whozthat? evolving an ecosystem for context-aware mobile social networks". In: *Network, IEEE* 22.4 (2008), pp. 50–55.

[129] Manoop Talasila, Reza Curtmola, and Cristian Borcea. "Improving location reliability in crowd sensed data with minimal efforts". In: *Wireless and Mobile Networking Conference (WMNC), 2013 6th Joint IFIP*. IEEE. 2013, pp. 1–8.

[130] Vigneshwaran Subbaraju, Amit Kumar, Vikrant Nandakumar, Sonali Batra, Salil Kanhere, Pradipta De, Vinayak Naik, Dipanjan Chakraborty, and Archan Misra. "ConferenceSense: monitoring of public events using phone sensors". In: *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM. 2013, pp. 1167–1174.

[131] Sanae Rosen, Sung-Ju Lee, Jeongkeun Lee, Paul Congdon, Z Morley Mao, and Ken Burden. "MCNet: Crowdsourcing wireless performance measurements through the eyes of mobile devices". In: *Communications Magazine, IEEE* 52.10 (2014), pp. 86–91.

[132] Yohan Chon, Nicholas D Lane, Yunjong Kim, Feng Zhao, and Hojung Cha. "Understanding the coverage and scalability of place-centric crowdsensing". In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM. 2013, pp. 3–12.

[133] Sasank Reddy, Deborah Estrin, and Mani Srivastava. "Recruitment framework for participatory sensing data collections". In: *Pervasive Computing*. Springer, 2010, pp. 138–155.

[134] Nicolas Haderer, Romain Rouvoy, and Lionel Seinturier. "A preliminary investigation of user incentives to leverage crowdsensing activities". In: *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on*. IEEE. 2013, pp. 199–204.

[135] Delphine Christin, Andreas Reinhardt, Salil S. Kanhere, and Matthias Hollick. "A survey on privacy in mobile participatory sensing applications". In: *Journal of Systems and Software* 84.11 (2011). Mobile Applications: Status and Trends, pp. 1928 –1946. ISSN: 0164-1212. DOI: http://dx.doi.org/10.1016/j.jss.2011.06.073. URL: http://www.sciencedirect.com/science/article/pii/S0164121211001701.

[136] Jeffrey A Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B Srivastava. "Participatory sensing". In: *Center for Embedded Network Sensing* (2006).

[137] Zhiwen Yu, Yun Feng, Huang Xu, and Xingshe Zhou. "Recommending travel packages based on mobile crowdsourced data". In: *Communications Magazine, IEEE* 52.8 (2014), pp. 56–62.

[138] David A Galvan, Attila Komjathy, Michael P Hickey, Philip Stephens, Jonathan Snively, Y Tony Song, Mark D Butala, and Anthony J Mannucci. "Ionospheric signatures of Tohoku-Oki tsunami of March 11, 2011: Model comparisons near the epicenter". In: *Radio Science* 47.4 (2012).

[139] Layla Pournajaf, Li Xiong, and Vaidy Sunderam. "Dynamic data driven crowd sensing task assignment". In: *Procedia Computer Science* 29 (2014), pp. 1314–1323.

[140] Xin Xiao, Xiaohua Tian, Xiaoying Gan, and Xinbing Wang. "Multi-Class Labeling with BCH codes for Mobile Crowdsensing". In: *Proceedings of IEEE GLOBECOM*. Austin, USA, Dec. 2014, pp. 265–270.

[141] Sanjit Kaul, Marco Gruteser, Vinuth Rai, and John Kenney. "Minimizing age of information in vehicular networks". In: *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on*. IEEE. 2011, pp. 350–358.

[142] Nikolaos Pappas, Johan Gunnarsson, Ludvig Kratz, Marios Kountouris, and Vangelis Angelakis. "Age of information of multiple sources with queue management". In: *IEEE ICC 2015*. London, United Kingdom, June 2015.

[143] B. Kantarci and H.T. Mouftah. "Trustworthy Sensing for Public Safety in Cloud-Centric Internet of Things". In: *Internet of Things Journal, IEEE* 1.4 (2014), pp. 360–368. ISSN: 2327-4662. DOI: 10.1109/JIOT.2014.2337886.

[144] Matthias Stevens and Ellie D'Hondt. "Crowdsourcing of pollution data using smartphones". In: *Workshop on Ubiquitous Crowdsourcing*. 2010.

[145] Yohan Chon, Nicholas D Lane, Fan Li, Hojung Cha, and Feng Zhao. "Automatically characterizing places with opportunistic crowdsensing using smartphones". In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM. 2012, pp. 481–490.

[146] *Amazon mechanical Turk: the online market place for work*. URL: https://www.mturk.com/mturk/welcome.

[147] Weisi Guo and Siyi Wang. "Mobile crowd-sensing wireless activity with measured interference power". In: *Wireless Communications Letters, IEEE* 2.5 (2013), pp. 539–542.

[148] Martin Azizyan, Ionut Constandache, and Romit Roy Choudhury. "SurroundSense: mobile phone localization via ambience fingerprinting". In: *Proceedings of the 15th annual international conference on Mobile computing and networking*. ACM. 2009, pp. 261–272.

[149] O. Yurur, C. Liu, Z. Sheng, V. Leung, W. Moreno, and K. Leung. "Context-Awareness for Mobile Sensing: A Survey and Future Directions". In: *IEEE Communications Surveys Tutorials* PP.99 (2014), pp. 1–1. ISSN: 1553-877X. DOI: 10.1109/COMST.2014.2381246.

[150] I. Carreras, D. Miorandi, A. Tamilin, E.R. Ssebaggala, and N. Conci. "Crowdsensing: Why context matters". In: *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*. 2013, pp. 368–371. DOI: 10.1109/PerComW.2013.6529518.

[151] Xiang Sheng, Jian Tang, Xuejie Xiao, and Guoliang Xue. "Sensing as a Service: Challenges, Solutions and Future Directions". In: *IEEE Sensors Journal* 13.10 (2013), pp. 3733–3741. ISSN: 1530-437X. DOI: 10.1109/JSEN.2013.2262677.

[152] B. Kantarci and H.T. Mouftah. "Sensing services in cloud-centric Internet of Things: A survey, taxonomy and challenges". In: *IEEE ICC Workshop (ICCW)*. 2015, pp. 1865–1870. DOI: 10.1109/ICCW.2015.7247452.

[153] S. Chatterjee, S. Misra, and S.U. Khan. "Optimal Data Center Scheduling for Quality of Service Management in Sensor-cloud". In: *IEEE Transactions on Cloud Computing* PP.99 (2015), pp. 1–1. ISSN: 2168-7161. DOI: 10.1109/TCC.2015.2487973.

[154] Charith Perera, Arkady Zaslavsky, Peter Christen, and Dimitrios Georgakopoulos. "Sensing as a service model for smart cities supported by Internet of Things". In: *Transactions on Emerging Telecommunications Technologies* 25.1 (2014), pp. 81–93. ISSN: 2161-3915. DOI: 10.1002/ett.2704. URL: http://dx.doi.org/10.1002/ett.2704.

[155] Yang Han, Yanmin Zhu, and Jiadi Yu. "A Distributed Utility-Maximizing Algorithm for Data Collection in Mobile Crowd Sensing". In: *Proc. IEEE GLOBECOM, Austin, USA* (2014).

[156] Nicholas D. Lane, Yohan Chon, Lin Zhou, Yongzhe Zhang, Fan Li, Dongwon Kim, Guanzhong Ding, Feng Zhao, and Hojung Cha. "Piggyback Crowd-Sensing (PCS): Energy Efficient Crowdsourcing of Mobile Sensor Data by Exploiting Smartphone App Opportunities". In: *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. SenSys '13. Roma, Italy: ACM, 2013, 7:1–7:14. ISBN: 978-1-4503-2027-6. DOI: 10.1145/2517351.2517372. URL: http://doi.acm.org/10.1145/2517351.2517372.

[157] Junxian Huang, Feng Qian, Alexandre Gerber, Z Morley Mao, Subhabrata Sen, and Oliver Spatscheck. "A close examination of performance and power characteristics of 4G LTE networks". In: *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM. 2012, pp. 225–238.

[158] C. Fiandrino, D. Kliazovich, P. Bouvry, and A.Y. Zomaya. "Network-assisted offloading for mobile cloud applications". In: *IEEE International Conference on Communications (ICC)*. 2015, pp. 5833–5838. DOI: 10.1109/ICC.2015.7249252.

[159] Andres Garcia-Saavedra, Pablo Serrano, Albert Banchs, and Giuseppe Bianchi. "Energy consumption anatomy of 802.11 devices and its implication on modeling and design". In: *Proceedings of the 8th international conference on Emerging networking experiments and technologies*. ACM. 2012, pp. 169–180.

[160] G. Foddis, R.G. Garroppo, S. Giordano, G. Procissi, S. Roma, and S. Topazzi. "LTE traffic analysis for signalling load and energy consumption trade-off in mobile networks". In: *IEEE International Conference on Communications (ICC)*. 2015, pp. 6005–6010. DOI: 10.1109/ICC.2015.7249279.

[161] A. Rahmati, C. Shepard, C. Tossell, Lin Zhong, and P. Kortum. "Practical Context Awareness: Measuring and Utilizing the Context Dependency of Mobile Usage". In: *IEEE Transactions on Mobile Computing* 14.9 (2015), pp. 1932–1946. ISSN: 1536-1233. DOI: 10.1109/TMC.2014.2365199.

[162] *DigiPoint 3: Points export from Google Maps*. 2014. URL: http://www.zonums.com/gmaps/digipoint.php.

[163] *OpenSignal*. 2016. URL: http://opensignal.com/.

[164] *Wi-Fi coverage Discover the vast extend of HOTCITY's Wi-Fi network!* 2016. URL: https://www.hotcity.lu/en/laptop/www/About/Wi-Fi-coverage.

[165] *Dangerous by design*. 2011. URL: http://www.smartgrowthamerica.org/documents/dangerous-by-design-2011.pdf.

[166] *The Ultimate Smart Thermostat*. 2015. URL: http://smartthermostatguide.com/.

[167] *foobot*. 2015. URL: http://foobot.io/features//.