

Why Energy Matters? Profiling Energy Consumption of Mobile Crowdsensing Data Collection Frameworks

Mattia Tomasoni^a, Andrea Capponi^{b,*}, Claudio Fiandrino^c, Dzmitry Kliazovich^d, Fabrizio Granelli^a,
Pascal Bouvry^b

^aUniversity of Trento, Trento, Italy

^bUniversity of Luxembourg, Luxembourg

^cIMDEA Networks Institute, Madrid, Spain

^dExaMotive, Luxembourg

Abstract

Mobile Crowdsensing (MCS) has emerged in the last years and has become one of the most prominent paradigms for urban sensing. The citizens actively participate in the sensing process by contributing data with their mobile devices. To produce data, citizens sustain costs, i.e., the energy consumed for sensing and reporting operations. Hence, devising energy efficient data collection frameworks (DCF) is essential to foster participation. In this work, we investigate from an energy-perspective the performance of different DCFs. Our methodology is as follows: (i) we developed an Android application that implements the DCFs, (ii) we profiled the energy and network performance with a power monitor and Wireshark, (iii) we included the obtained traces into CrowdSenSim simulator for large-scale evaluations in city-wide scenarios such as Luxembourg City, Turin and Washington DC. The amount of collected data, energy consumption and fairness are the performance indexes evaluated. The results unveil that DCFs with continuous data reporting are more energy-efficient and fair than DCFs with probabilistic reporting. The latter exhibit high variability of energy consumption, i.e., to produce the same amount of data, the associated energy cost of different users can vary significantly.

Keywords: Mobile Crowdsensing, Energy Consumption, Data Collection

1. Introduction

Nowadays, half of the global population lives in urban areas. This figure is projected to further increase in the next three decades [1]. The unprecedented population growth calls for a sustainable development of urban areas. Currently, significant research efforts are undergoing to provide efficient services exploiting ICT solutions, such as the Internet of Things (IoT) paradigm. Smart cities aim to

*Corresponding author. Tel.: (+352) 46 66 44 5320

Email addresses: mattia.tomasoni@studenti.unitn.it (Mattia Tomasoni), andrea.capponi@uni.lu (Andrea Capponi), claudio.fiandrino@imdea.org (Claudio Fiandrino), kliazovich@ieee.org (Dzmitry Kliazovich), fabrizio.granelli@unitn.it (Fabrizio Granelli), pascal.bouvry@uni.lu (Pascal Bouvry)

improve the quality of life of citizens¹ by providing innovative solutions for public services, such as health care, urban safety, waste and emergency situation management [2]. Sensing is crucial in monitoring infrastructures, transportation systems, environment, health but also resource consumption [3, 4]. While deploying and maintaining sensing infrastructures is very expensive, including citizens in the
10 loop of sensing is cost-effective and was proved to be a win-win strategy. The citizens' participation to sensing activities is at the heart of the Mobile Crowdsensing (MCS) paradigm [5].

MCS has become an appealing paradigm for urban sensing in the recent years. In MCS, citizens contribute data produced by sensors embedded in their devices, such as smartphones, tablets and wearables [6]. The key idea of MCS is to integrate human intelligence with machine intelligence [7].
15 Active participation of citizens guarantees higher coverage and deeper context awareness compared to traditional sensor networks. Furthermore, users ensure self-maintenance and recharge of the devices that act as sensor and communication nodes. MCS systems fit well as large-scale sensing platforms for several types of applications such as emergency situation management, intelligent transportation systems and health care [8]. A particular example is environmental monitoring such as detection of
20 obstacles [9] or potholes [10], and monitoring of watershed conditions [11]. Garbagewatch [12] relies on citizen participation to improve the collection and recycling process, while the National Environment Agency of Singapore employs on a daily basis HazeWatch² to perform air pollution monitoring, an application previously developed at The University of New South Wales [13]. In all such applications, the collected data is sent through cellular (e.g., 4G/LTE) or WiFi data connection to a central collector,
25 which is usually located in the cloud [14]. The organizer is usually responsible for user recruitment and task assignment prior to the start of the sensing campaign and for data analysis and processing during the sensing campaign. The organization of a MCS campaign requires to sustain costs to reward individuals for their involvement and to verify the accomplishment of the tasks. Consequently, it is crucial to investigate how to maximize the efficiency of a Data Collection Frameworks (DCF), which is
30 defined in terms of the costs sustained by the organizer and the revenues [15]. The users sustain costs while contributing data too, i.e., the energy spent from the batteries for sensing and reporting data and, eventually, the data subscription plan if cellular connectivity is used for reporting. Developing efficient DFCs is crucial to regulate the degree of user involvement and to prevent excessive battery drain from the mobile devices, which are fundamental limiting factors to foster user participation and
35 contribution [16]. At the same time, a DCF has to gather a sufficient amount of data to ensure Quality of Information (QoI) [17, 18].

This paper takes one step further the problem of energy efficient data collection in MCS systems.

¹In the remaining part of the paper, we use terms citizens, crowd, participants and users interchangeably.

²Available at: <https://www.haze.gov.sg/> and <https://play.google.com/store/apps/details?id=com.vulcan.hazemon>

Despite the scientific community made numerous attempts to build effective large-scale MCS applications, the design of efficient DCFs remains a grand challenge and practically few works have attempted to study from an experimental perspective the performance of DCFs. With this work³, we compare the performance of different DCFs by employing an experimental methodology. We consider DCFs that differ by the following features: (i) the type of data reporting⁴ mechanism (DRM) implemented, e.g., continuous (i.e., data is delivered while produced), delayed (i.e., data is delivered once the sensing activity has ended) and probabilistic (i.e., data transmission is occasional during sensing), and (ii) the degree of control the collector establishes through feedback on the amount of data is still to be harvested. All these aspects fundamentally impact the efficiency of DCFs. We developed an Android application that implements the various DCFs and performed energy measurements on the smartphone through a power monitor. Furthermore, we also collected network-related measurements exploiting Wireshark, a free and open source packet analyzer. As MCS systems require large participation to be effective, performing feasible experiments on real testbeds is often not feasible. To this end, simulations are a valid alternative. Hence, the collected energy- and network-measurements were employed to feed CrowdSenSim simulator for large-scale evaluations in city-wide scenarios. CrowdSenSim [20] is the first tool to perform scalable MCS activities for a required duration (e.g., days) in complex environments.

The synopsis of contributions of this work is as follows:

- Analysis and comparison of DCFs with different DRMs, i.e., continuous, delayed and probabilistic data delivery.
- Implementation of an flexible Android application that supports the aforementioned DRMs to experimentally profile: (i) the energy costs associated to sensing and reporting operations, and (ii) communication features of data reporting such as the packet send rate and the distribution of packet errors.
- Inclusion of the DRMs energy profiles obtained from the analysis with the power monitor and implementation of different DCFs in CrowdSenSim simulator for large-scale performance evaluation in urban scenarios.

Our main findings are as follows:

- The type of DRM is the key criterion that differentiates the DCFs. DCFs with continuous reporting that implement mechanisms to block sensing and data delivery based on history of user contribution are more effective in harvesting data from the crowd.

³A preliminary version of this work appeared in [19]

⁴In the remaining part of the paper, we use terms data reporting and data delivery interchangeably.

- As expected, DCFs with probabilistic reporting exhibit high variability of energy consumption, i.e., to produce the same amount of data, the associated energy cost of different users can be significantly different.
- Human mobility does not influence the per-user energy-consumption behavior of the DCFs. Experiments performed on cities with different urban morphology show that the variation of the average per-user energy consumption achieved with the various DCFs is minor.

The remainder of the paper is organized as follows. Section 2 presents background on DCFs in MCS and motivates the need for efficient DCFs. Section 3 explains how to experimentally profile performance through real measurements on a smartphone. Section 4 proposes our novel methodology to assess the performance of DCFs through simulations on a large scale. Section 5 illustrates performance evaluation conducted with the CrowdSenSim simulator and, Section 6 provides a conceptual analysis of networking changes in crowdsensing applications with the advent of the fifth generation (5G) of cellular mobile networks. Finally, Section 7 concludes the work presenting future directions.

2. Background and Motivation

This Section overviews related works on data collection in MCS (Subsection 2.1) and presents the three DCFs under analysis (Subsection 2.2).

2.1. Related Works

DCFs are developed to support data collection useful to many applications at the same time. These DCFs usually aim to maximize a set of parameters, e.g., the temporal and spatial coverage of the contributed data or the QoI while minimizing at the same time the costs, such as energy consumption or monetary rewards [21]. Wu et al. [22] investigate the trade-off between the amount of acquired data and the associated energy consumption. The authors present and analyze both off-line and on-line settings for task allocation. In off-line case, the entire task information is known a-priori and does not change over time, while in the on-line scenario tasks are dynamically allocated in real-time without any information in advance. The authors first provide an optimal algorithm for the off-line setting. Then, they investigate the on-line setting where requests arrive dynamically without prior information in a first-in-first-out (FIFO) manner or with an arbitrary deadline (AD). Wang et al. [23] investigate the problem of scheduling multiple sensing tasks with the objective of ensuring the quality of sensed data while minimizing the energy consumption. Starting from basic cases in which sensing process requires data from only one sensor, the authors define the Minimum Energy Single-sensor task Scheduling (MESS) problem and design a polynomial-time optimal algorithm. Then, they consider a generic case in which sensing tasks require data from multiple sensors to be accomplished. To solve the

problem of Minimum Energy Multi-sensor task Scheduling (MEMS), the authors propose an Integer Linear Programming (ILP) formulation as well as two effective polynomial-time heuristic algorithms. In [24], the authors propose a fair energy-efficient allocation framework whose objective is to minimize the maximum aggregated sensing time. The problem is NP-Hard also when the information on the tasks such as arrival and duration is known prior to the allocation. The authors first investigate the off-line allocation model and propose an efficient polynomial-time approximation algorithm with a factor of $2 - 1/m$, where m is the number of mobile devices joining the system. Then, focusing on the on-line allocation model, they design a greedy algorithm which achieves a ratio of at most m . Liu et al. [25] propose a method to efficiently select users for participatory crowdsensing. Contributors are dynamically chosen considering their willingness to acquire data and their potential, which is calculated considering the remaining battery in their smartphones. The distribution of tasks aims to minimize the probability that an individual does not accomplish the assigned task. The CARDAP [26] DCF exploits a fog computing platform to enable efficient data analytics performed in a distributed fashion. The fog allows CARDAP to extend and augment functionality of a previously proposed general-purpose framework called CAROMM [27]. Similarly to CARDAP, the framework proposed in [16] exploits the fog to perform user recruitment based on multiple criteria, including distance of the participants from the location of the sensing task, their remaining battery charge and the user sociability defined in terms of the amount of time and data users exchange through social media. Wang et al. [28] present an algorithm to report information in an energy-efficient way. It classifies users into two groups. In the first category, the target is to minimize the energy consumption while reporting data and the individuals pay for the data they utilize to the operators. In the second group, users aim to minimize the cost of data reporting using communication technologies such as WiFi or Bluetooth, which are free-of-charge.

To the best of our knowledge, there are no existing studies in the literature that investigated and compared the performance of DCFs for large scale sensing campaigns to mimic a real MCS deployment. In this work, we tackle the problem from a particular angle, by first implementing and measuring the performance of DCFs and then performing large-scale simulations with CrowdSenSim. The closest to our work is [29], where Peng et al. investigate the trade-off between energy efficiency and fairness between participants. They propose an allocation framework which aims to minimize the maximum aggregate sensing time and for the experimental setup they exploit the power monitor to assess the power consumption of a smartphone under different configurations.

2.2. DCFs under Analysis

This work compares the performance of three different DCFs. Each of them has specific properties and features that are presented in the following paragraphs. Other existing DCFs in the literature can

135 be categorized as intermediate solutions of the studied ones.

DDF - Deterministic Distributed Framework. DDF is a DCF with the objective of fostering energy-efficient data collection in opportunistic cloud-based MCS systems that we proposed in a former work [30]. It aims at maximizing the utility of the cloud collector in receiving data from certain sensors in a specific region of interest, while minimizing at the same time the energy costs users sustain to
140 sense and deliver information. The central collector sends periodically beacons to the mobile devices to share the utility in receiving data from specific sensors. Then, the mobile devices take sensing decisions locally and independently one with each other. Sensing and reporting decision are driven by several factors such as the environmental context, the potential cost for sensing and reporting, the current level of battery and the amount of data already contributed. When these factors are above predefined
145 thresholds, the systems prevent users from further contribution. Therefore, this mechanism is aware of the *user-state*, i.e., it considers the user history to drive the subsequent decisions. The DCF aims to be fair in terms of the amount of energy spent by the users.

PDA - Probabilistic Distributed Algorithm. Montori et al. propose a distributed algorithm based on probabilistic design to acquire data in an opportunistic scenario [31]. The algorithm is based
150 on limited feedback from the central collector and does not require users to complete specific task. The objective of this algorithm is to regulate the amount of data contributed from users in a certain region of interest to minimize data redundancy and energy waste. Assuming that it is impossible to compute the number of participants in a region of interest because their position can not be disclosed for privacy reasons, the coordinator determines the number of participants through the number of observations
155 received. The ultimate objective is to control the number of observations in an area in order to guarantee a desired value, i.e., the minimum quantity of data to be harvested. To reach this goal, the mobile devices decide independently from the collector and one with each other when to perform sensing and reporting. The framework is memoryless. Unlike DDF, it does not take into account the previous status of each user to determine transmission opportunities, which are independent events.
160 In addition, the proposed algorithm aims at maximizing the fairness among participants in order to not have few devices contributing much data.

PCS - Piggyback Crowdsensing. PCS is a DCF that performs data reporting exploiting the so called smartphones opportunities [32]. The authors define *smartphone opportunities* those times when smartphone users place phone calls or use applications. To this end, PCS significantly lower energy
165 consumption associated to reporting operations. Indeed, during smartphones opportunities, the mobile devices do not have to wake up the radio interface to transmit the collected data on purpose. PCS does not include any feedback from the collector to the users, hence it can not indicate urgency for additional data. Furthermore, if smartphones opportunities are rare, the collector might suffer because

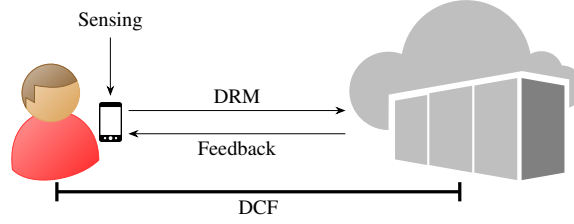


Figure 1: MCS data collection end-to-end: role of DCF and DRM from the user to the cloud collector

of missing data. According to these considerations, PCS is particularly suitable for delay-tolerant MCS
 170 tasks that do not need data to be sent to the central collector as soon as users sense it.

3. Profiling Energy Consumption of DCFs

This Section presents the methodology employed to experimentally assess the performance of DCFs. We developed a custom Android application for smartphone which implements the aforementioned DCFs. We perform energy measurements with a power monitor while network-related measures are
 175 obtained with Wireshark. The results obtained and presented in this Section will be exploited in next Section to assess the performance of the DCFs with simulations on large scale with CrowdSenSim.

3.1. Information Delivery: the Data Reporting Mechanisms (DRM)

A DRM is the most important component of a DCF. It defines the methodology to perform delivery of sensed information to the cloud collector. Note that a DCF consists of multiple components in
 180 addition to the DRM, such as mechanisms to inform the users about the urgency of sensing additional data or to prevent transmission of not-any-longer needed information (e.g., through feedback from the collector) and the policy to perform sensing. Figure 1 graphically shows the components of a DCF. Every DCF implements a DRM: the ones presented in the following paragraphs are three methods that represent popular techniques for data delivery.

185 **Continuous-DRM (CON)** Data can be delivered in a continuous fashion as soon as it is sensed. This approach is needed for real-time applications where users need to feed the collector with data constantly over time. However, such continuous stream of data incurs in the highest energy cost from the user point of view as the mobile devices need to maintain the connection active during the entire sensing operation period. The CON method is implemented by DDF.

190 **Delayed-DRM (DEL)** With delayed reporting, data is sent after the sensing activity has ended. Hence, this approach decouples sensing and reporting operations and is more conservative than CON from an energy-consumption perspective. Indeed, the network interfaces do not have to be maintained continuously active for a prolonged period of time. DEL is useful for delay-tolerant applications where

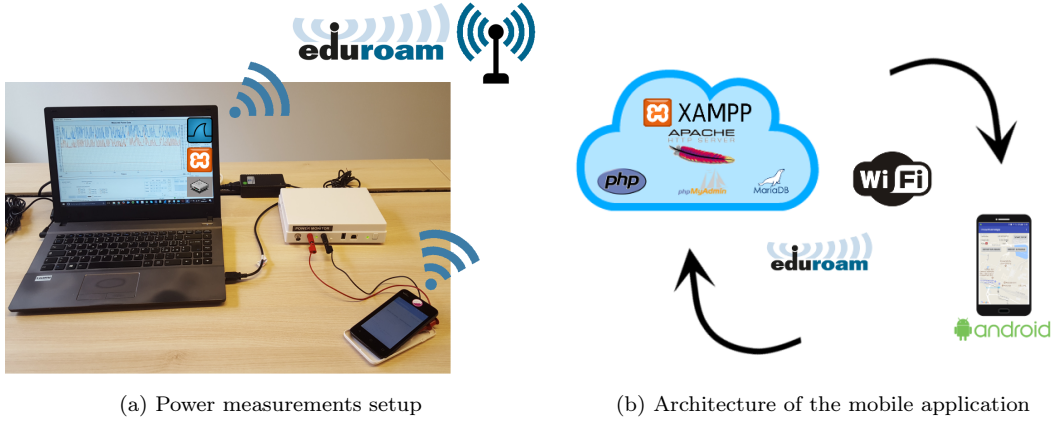


Figure 2: Experimental set-up

the collector does not need to harvest data in real-time. DEL is implemented by PCS, which exploits
 195 phone calls and application usage as smartphone opportunities.

Probabilistic-DRM (PRO) This approach considers a probabilistic reporting. It is an intermediate solution between the two previous approaches. During each timeslot, the algorithm randomly generates a probability p that drives the decision of delivering the sensed data. The parameter p is checked against δ , a feedback provided by the collector. This DRM is implemented by PDA .

200 3.2. The Application Architecture in a Nutshell

The Android mobile application follows the REST guidelines and is flexible as it implements and can be configured to operate with any of the DRMs presented in Subsection 3.1. Before describing the architecture in detail, the following paragraph briefly overviews important existing works and standard approaches to implement MCS applications.

205 3.2.1. Background

APISENSE [33] is a popular cloud-based platform that enables researchers to deploy crowdsensing applications by providing resources to store and process data acquired from a crowd. It presents a modular service-oriented architecture on the server-side infrastructure that allows researchers to customize and describe requirements of experiments through a scripting language. In addition, it
 210 makes available to the users other services (e.g., data visualization) and a mobile application, allowing them to download the tasks, to execute them in a dedicated sandbox and to automatically upload data on the server. In [34], Paganelli et al. propose a framework to support developers in modeling web resources, exposing them through RESTful APIs and developing applications. The framework consists of a Web Resource information model, a middleware, and tools for developing and publishing smart
 215 things' digital representations on the Web. They discuss the framework in compliance with REST

guidelines and its major implementation choices. Mulero et al. [35] present an infrastructure for smart cities that combines IoT and Linked Open Data paradigms to provide a scalable and responsive system able to provide services. They exploit a REST API that receives and manages a large amount of data. CRATER [36] is a crowdsensing platform to estimate road conditions. It provides RESTful APIs to
220 access data and visualize maps in the related application.

Similarly to previous works in the area, we followed the REST guidelines to develop our application. The REST guidelines define the methodology to develop web services. In brief, for an application to be REST-compliant, the resources have to be self-descriptive and linked one with each other. It requires an uniform interface to decouple the architecture and enable each part to evolve independently, exploiting
225 basic HTTP methods (e.g., GET and POST to exchange messages and data with the cloud) and HTTP response. Furthermore, our custom application has a client-server architecture and stateless communications because each message is self-descriptive through HTTP headers. In other words, to minimize the server’s workload each request has enough information for the server to process that message.

230 3.2.2. Our Application

The mobile application we developed can run over any Android-based smartphone. Fig. 2(b) illustrates the architecture with all its components. The figure also highlights the public WiFi network EDUROAM that provides interconnection between the smartphone and collector. Java is the programming language employed for the implementation, while PHP is the server-side scripting language
235 used for the web development. The minimum supported version is Android Marshmallow 6.0 (API level 23), but the application is already compatible with Android Oreo 8.0 (API level 26). The server side, i.e., the cloud collector is a laptop used to perform data processing and storage. In particular the system exploits XAMPP (v7.1.8 - 32bit), that provides in a unique distribution Apache web server and phpMyAdmin to manage the database. The most recent version of XAMPP features a database
240 based on MariaDB.

The unique features of each DCF impact the DRM implementation design. For example, DEL requires a database to locally store the sensed data. CON and PRO are less strict and will work correctly if the sensed data is simply stored on a local buffer. A database ensures higher reliability as it adds a further level of protection at application layer on top of communication retransmissions
245 schemes at lower layers of the protocol stack.

3.3. Experimental Set-up

Energy profiles of these policies are obtained through the utilization of a power monitor, similarly to previous research [37, 38]. Fig. 2(a) shows the setup for the measurements performed with the

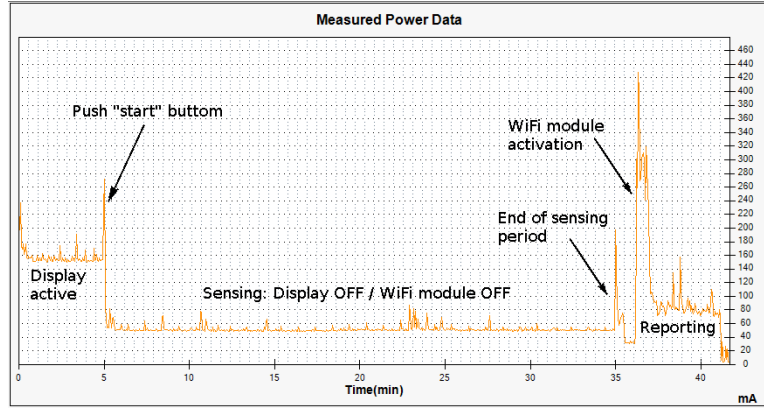


Figure 3: Screenshot of power monitor measurements

smartphone under analysis, which is a Wiko Sunny⁵. In the experiments, the smartphone runs Android
 250 Marshmallow version 6.0 (API Level 23) and is equipped with a quad-core 1.3 GHz Cortex-A7 processor
 and has 512 MB of RAM. The smartphone also features 5 MP camera and 8 GB of flash storage. The
 smartphone provides WiFi and Bluetooth connectivity, supporting 802.11 b/g/n and Bluetooth v4.0
 standards respectively. It is powered by a 1200 mA, 3.7 V battery. In the experiments, the application
 performs sensing from GPS, accelerometer and proximity sensors of the smartphone. The smartphone
 255 delivers data to a laptop that also run the software to analyze the power measurements. The laptop
 acts like the cloud collector. Since the computational power and the hardware of the laptop offer
 considerably exceeding performance of the mobile device, the hypothesis is consistent. The laptop is
 a equipped with a dual-core 2.6 GHz Intel i5-4210M and 8 GB of RAM; it has a 256 GB Crucial SSD
 as storage and an Realtek card for WiFi 802.11 b/g/n connectivity. The laptop runs Windows 10
 260 OS at the time of the tests. To accomplish the power measurement campaign, we utilize the power
 monitor hardware by Monsoon⁶. Existing works acquired power consumption measures via software
 by means of system calls [39, 40]. We exploit a different method because the power monitor directly
 retrieve measurements while ensuring a higher level of accuracy. In order to collect data, the power
 monitor needs to power the smartphone directly, hence in the equivalent circuit it substitutes the
 265 internal battery. 5000 samples per second is the sampling rate to record measures in real-time. Users
 can export readings in a csv file at the end of the measure campaign exploiting a specific software,
 which also displays a real-time chart of the measurements.

⁵ Available at: <http://it.wikomobile.com/m1330-sunny>

⁶ Available at: <http://www.monsoon.com/LabEquipment/PowerMonitor/>

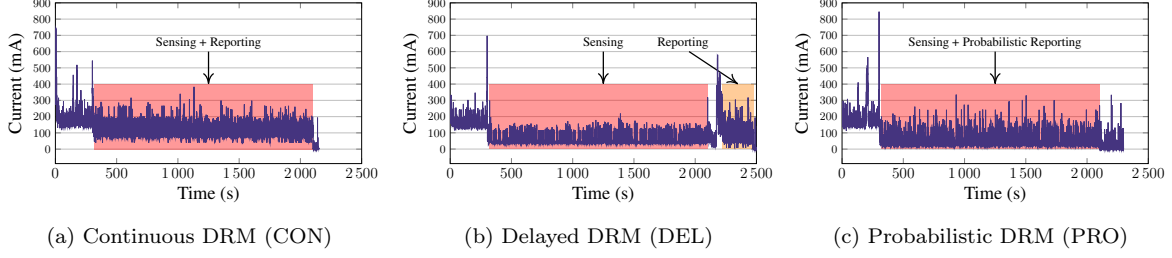


Figure 4: Real energy measurements performed with power monitor for different data reporting mechanisms (DRMs)

3.4. Experimental Results

This Section presents the results obtained from energy and network measurements performed with the power monitor and Wireshark for the DRM. The challenge in performing the energy measurements is to find the right baseline for comparison. Specifically, different DRMs spend different times for reporting, i.e., CON keeps active the interface all the time, while with DEL, the time in which the interface remains active is supposed to be shorter. For such a reason, we base our measurement campaign considering for all DRMs the same amount of time spent for sensing, i.e., 30 min. Consequently, all the DRMs gather the same amount of information that needs to be delivered. However, each DRM spends different amounts of time for data reporting. For example, DEL performs sensing for 30 min and then it takes other 6 min to perform reporting.

Fig. 3 shows an example of measurement profile of the application for DEL and highlights the various activities in each point in time. We test the PRO mechanism with a feedback from the collector δ set to 0.5 for comparison with other schemes DEL and CON, but we perform additional tests with various settings, namely $\delta = [0.25, 0.5, 0.75]$. We performed the measurements with the EDUROAM public WiFi network. This allows to test the performance of the DRMs in realistic environments while setting-up an ad-hoc network for the tests would have achieved better results as the environment is totally under control.

Fig. 4 compares the instantaneous current drain of the proposed DRMs while reporting data. In the continuous reporting approach (CON), the flow of data to transmit is bursty because the mobile device exploits transmission opportunities as sensing readings are acquired (see Fig. 4(a)). Consequently, the transmission time is short, but at the price of higher current drain to maintain the interface active. Conversely, the delayed reporting approach (DEL) makes the transmission period shorter because data transmission occurs at the end of the sensing process, allowing to establish and terminate the connection in a small amount of time. Fig. 4(b) clearly shows that the current drain attributed to sensing process is on average lower than CON method, then the activation of the interface generates a peak around 2150 s. The probabilistic reporting approach (PRO) provides intermediate performance with respect to the two previous approaches (see Fig. 4(c)). The network interface remains active for

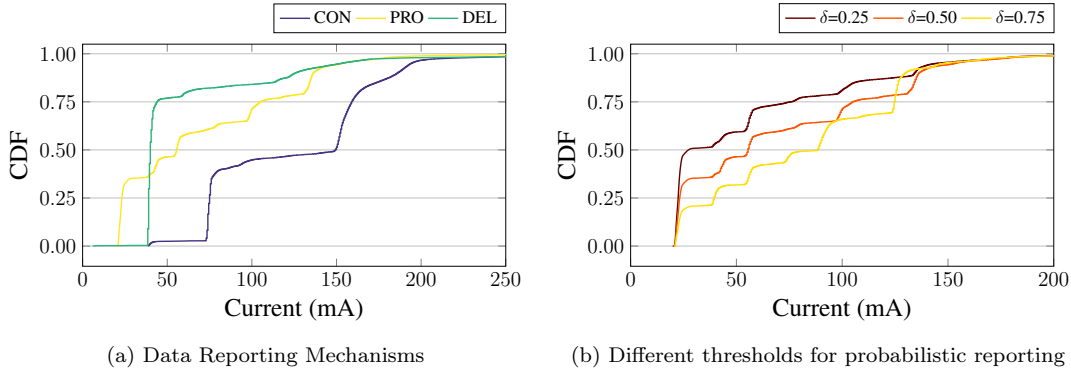


Figure 5: CDF of energy spent for different reporting approaches

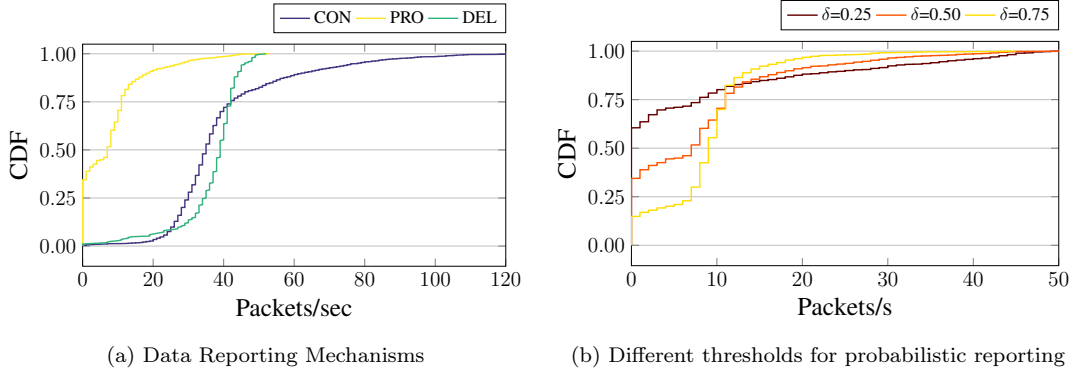


Figure 6: Distribution of transmission rates

295 a prolonged period of time, longer than CON, but with an on-average lower current drain.

Fig. 5 shows the CDF of the battery drain. In particular, Fig. 5(a) focuses on the various DRMs while Fig. 5(b) focuses on the PRO mechanism and highlights the impact of δ , the probability of transmission in each timeslot. Note that each timeslots is worth 40 s. With CON, the battery drain is low than 75 mA for a significant fraction of the time. We highlight that such value of current is the peak value measured in a given point in time. Compared to CON, DEL exhibits on average with higher instantaneous peak values. During 50% of the reporting time, while CON exhibits peak values above 150 mA, DEL achieves values above 40 mA. Such behaviour is expected from the theoretical results on WiFi energy consumption as one of the substantial components to the total energy budget depends on the traffic load [41]. On the one hand, CON maintains the interface active for longer but transmitting few packets each time. On the other hand, DEL sends higher bursts of packets during shorter time periods. PRO exhibits an intermediate behaviour than CON and DEL. When analyzing PRO in Fig. 5(b), the higher the value of δ , the higher the probability to transmit. This translates in a different distribution of the instantaneous peak values of current: with $\delta = 0.75$, the peak values are below 75 mA for 75% of the reporting time, with $\delta = 0.5$ this figure becomes 62.5% of the reporting time and drops to 48% for $\delta = 0.25$.

310

Fig. 6 shows the CDF of the packet rate transmission obtained with Wireshark. Fig. 6(a) compares

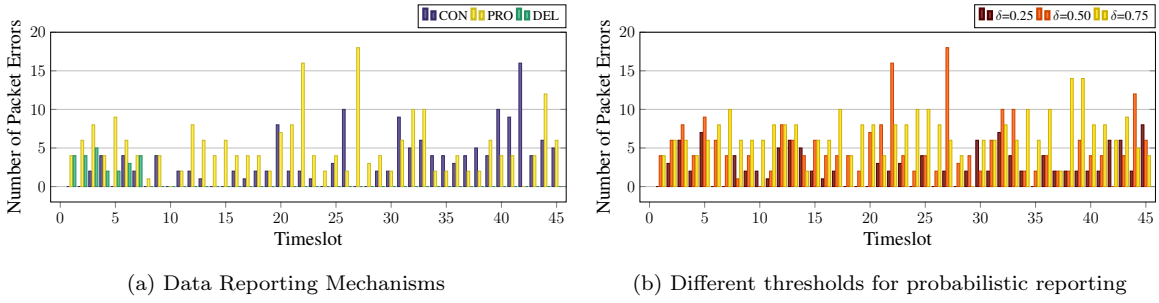


Figure 7: Distribution of packet errors

the DRMs that exhibit substantially different distributions of packet transmission rates. Indeed, for 75% of the reporting time, PRO achieves rates as high as 10 packets/s. Compared to PRO, both CON and DEL achieve rates as high as 40 packets/s for 75% of the reporting time. Interestingly, unlike PRO and DEL that converge to a certain maximum rate, CON has higher variability and packet rates can be as high as 120 packets/s. The reason is due to the technical implementation. While DEL is basically transmitting a unique (and bigger) file, CON transmits frequently smaller chunks. Hence, the collector and the mobile phone have to synchronize much more often to ensure reliable data delivery. Similarly to the result obtained for the energy (see Fig. 5(b)), Fig. 6(b) shows that also the transmission rate varies with the increase of δ . Interestingly, the variation is evident only for low rates. For example, rates up to 5 packets/s occurs for 20%, 40% and 70% of the reporting time for $\delta = 0.25$, $\delta = 0.5$ and $\delta = 0.75$ respectively.

Fig. 7 shows the distribution of packet errors along time. We recall that each timeslot is 40 s long and that a public WiFi network provides interconnection between the collector and the smartphone. Fig. 7(a) compares the DRMs. As expected, the distribution of packet errors for DEL is concentrated as the data deliver is shorter. CON and PRO DRMs exhibit high variability in packet errors that is due to the realistic environment as that the public network is not under our control. However, it should be noted that PRO experiences a comparatively higher number of losses. Indeed, the 802.11 protocol is well known to be inefficient as its scheduling strategy is to allocate single resources to single nodes [42]. Hence, it favors CON rather than PRO types of data transmissions.

4. Large-scale Analysis: the Methodology

This Section presents the methodology employed to perform the large-scale evaluation. In particular, Subsection 4.1 introduces the platform for simulations and Subsection 4.2 presents the performance metrics.

335 4.1. A Primer on CrowdSenSim

CrowdSenSim is a custom simulator that we have developed to assess the performance of crowd-sensing activities in large urban areas [20]. CrowdSenSim supports pedestrian mobility in citywide scenarios. All the individual walking paths are known before simulation runtime to ensure scalability of the platform. Users contribute data according to the implemented DCF during runtime and the simulator computes the amount of gathered data for each user and the associated battery drain for
340 sensing and reporting [30]. Data generation exploits sensors that are typically available in mobile devices. For this work, the details of the smartphone described in Section 3.3. The simulation results can be obtained at the level of individual devices as well as the system level, which helps to analyze DRMs and how much employed crowdsensing techniques are effective. The simulator is built with
345 independent modules can inter-operate one with each other. The simulator allows to perform different crowdsensing-related analysis such as user rewarding, task allocation or fairness between participants. We have implemented PDA and PCS DCFs to assess and compare them with DDF. The implementation of the DCFs is based on the energy consumption model previously described and the detailed description of amount of contributed data and battery drain is explained in detail in Subsection 5.1.
350 Human mobility is based on the ParticipAct dataset that was generated by a real MCS campaign of approximately 170 students in the Emilia Romagna region (Italy) [43]. To establish the user arrival pattern in the simulator, we extracted the profile of the average number of contacts during 7 days without having at disposal the dataset. Specifically, we divide into hours the total simulation period estimating the minimum number of individuals to be allocated to have an average user contact fol-
355 lowing the ParticipAct profile. By definition, a unique user contact is the overlap within a timeslot of two user walking paths so that their distance is below a radius R . For an implementation choice, multiple overlaps in different timeslots count as distinguished contacts, while multiple overlaps in the same timeslot are considered unique contacts.

To assess the energy consumption for a sensing campaign in real urban environments, we include
360 in CrowdSenSim real traces of energy measurements for sensing and reporting. Specifically, the cost that each device experiences is computed proportionally to the time of contribution. The reference power consumption profiles were obtained from 30 min long sensing traces (see Fig. 5(a)).

4.2. Performance Metrics

Amount of contributed data. The main objective of any DCF is to gather a sufficient amount of
365 information to capture and or monitor phenomena. Hence, the quantification of the amount of data that the crowd can harvest is important to assess. Other metrics related to QoI and accuracy of data are equally important, but in this work we focus solely on the amount of contributed data as the main objective is to compare the DCFs.

Energy consumption. This performance metric defines the costs that users mobile devices experience
 370 and is measured in mAh.

Fairness. Ideally, MCS systems should gather information from the crowd by ensuring QoI. At the same time, the collector should guarantee fair treatment to each of the participants, i.e., it should not take advantage from a small set of users. Intuitively, users that contribute higher amounts of data, sustain a higher energy cost to produce such data and therefore need to be rewarded adequately.

375 To evaluate fairness between users in a DCF, we exploit the Jain Fairness Index [44]. The index measures the equality in allocating a set of finite resources to users according to certain criteria. For instance, if a user is willing to pay a sum that is twice more high than other users for a shared resource, the scheduling algorithm should allocate to him twice the amount of resources than the other users. In MCS context, we interpret this index as a way to measure the fairness in which a set of users contribute
 380 to create a finite set of resources, i.e., the total amount of data delivered to the collector. Specifically, we measure the degree of equality in which each DCFs make users to produce their individual amount of data and the associated energy cost. First, we define separate indexes for each of the two components and then we define a global fairness index to incorporate both components.

Intuitively, users that walk for longer time periods are expected to contribute more data than others. We define the data contribution fairness index (F_D) as follows:

$$F_D = \frac{\left(\sum_{i=1}^N d_i\right)^2}{N \cdot \sum_{i=1}^N d_i^2}, \quad (1)$$

where:

$$d_i = \frac{D_i}{D_i^M}. \quad (2)$$

D_i is the amount of contributed data from user i and D_i^M is the maximum amount of information an individual could contribute in the corresponding time. F_D assumes values equal to 1 when users contribute data proportionally to the time spent walking. However, this index presents a significant drawback if considered alone. Indeed, F_D does not distinguish between two users that walk for the same time period, but have different initial levels of battery. For this reason, we introduce an additional index that takes into account the battery level of the devices. Specifically, a device with a higher battery level prior to the start of the sensing process is expected to contribute higher amounts of data than devices with a lower battery level. The index of battery fairness index (F_B) is defined as:

$$F_B = \frac{\left(\sum_{i=1}^N b_i\right)^2}{N \cdot \sum_{i=1}^N b_i^2}, \quad (3)$$

where:

$$b_i = \frac{B_i}{B_i^T}. \quad (4)$$

B_i and B_i^T are measurements of battery level of the mobile device i (in mAh). The former is the amount of *battery drain* experienced during the contribution process and the latter is the total battery when the mobile device starts to contribute data.

We introduce the crowdsensing fairness index (F_{CS}) to simultaneously take into account both the battery drain and the amount of contributed data. Specifically:

$$F_{CS} = \sigma \cdot F_D + (1 - \sigma) \cdot F_B, \quad (5)$$

where σ is a balancing coefficient that assumes real values in $[0, 1]$ and weights the relative importance between the two indexes F_D and F_B .

5. Performance Evaluation

This Section at first introduces the evaluation scenario (Subsection 5.1) by explaining in detail the simulation setting and then presents the results (Subsection 5.2).

5.1. Simulation Setting

In CrowdSenSim, the layout of the city is defined in terms of a set of coordinates \mathcal{C} containing information on $\langle \text{latitude}, \text{longitude}, \text{altitude} \rangle$. The set of coordinates compose the street network of the urban area where the participants will move during runtime. They are obtained through the combination between an algorithm we developed running in the background of CrowdSenSim [45] and OSMnx [46]. OSMnx is a Python package that allows the researchers to easily download and analyze street networks of any location in the world from OSM, which is a map created by people under an open license. It simplifies the inconsistent node topology of OSM, but it does not provide a sufficient level of fine-grained details to MCS purposes. To fill this gap, our algorithm augments the resolution of street nodes, providing the possibility to simulate participants' mobility and task allocation with higher precision.

Fig. 8 shows the selected cities where users can move during simulation runtime: Luxembourg City (see Fig. 8(a)), Turin (Italy) (see Fig. 8(b)), and Washington DC (USA) (see Fig. 8(c)). The rationale for the choice of Luxembourg City, Turin and Washington DC is twofold. First, the selected cities can be ordered by growing size. The center of Luxembourg City covers an area of 51.47 km² with a population of 114 090 inhabitants as of the end of 2016 and is a fast growing city with headquarters of many international institutions (see Fig. 8(a)). The city center of Turin occupies an area of 130.17 km² and has a population of 883 601 inhabitants as of the beginning of 2016 (see Fig. 8(b)). The city center of Washington DC covers approximately an area of 158.1 km² with a resident population of 672 228

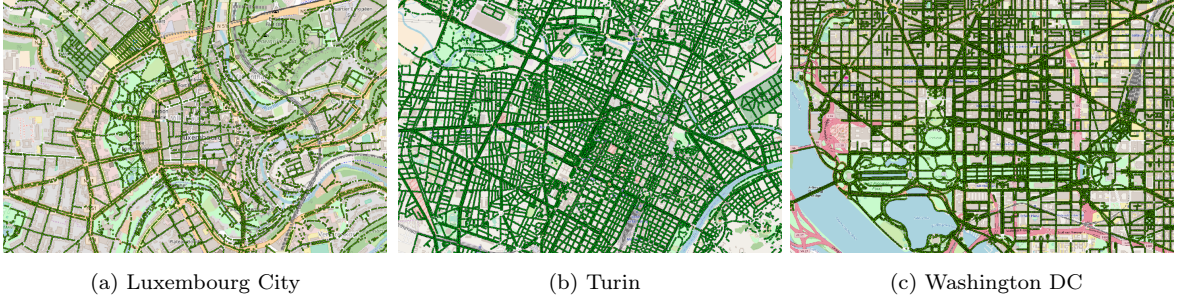


Figure 8: Cities taken into account for the evaluation. The green dots represent the pedestrian lanes where users walk obtained with an algorithm running in the background of CrowdSenSim

inhabitants as of the end of 2015 (see Fig. 8(c)). The second motivation behind the choice is about the urban morphology, which determines the street network topology. Luxembourg City shows the common north european urban morphology with many short streets with small lanes, a high density of crossroads in the center and few parallel large streets in the periphery. Washington DC is totally different from Luxembourg City and its topology of street network presents large lanes with a high number of parallel long streets. In addition, the differences between the urban morphology in the city center and the periphery are minimal. Turin falls in between the two former categories because of typical roman grid street organization.

As explained in Section 4.1, the user arrival pattern exploited in CrowdSenSim is based on realistic mobility traces and chosen simulation period is 12 consecutive hours in one day. The ParticipAct dataset supplies information on the user contact per-hour. Following this idea, the simulator assigns a certain number of participants to arrive at the desired sum of contacts for each hour. In detail, we consider a user contact when two walking paths are within a certain timeslot, so that two individuals are in a radius of 50 m. Only one mobile device for each user generates data. The users walk with an average speed uniformly distributed between $[1, 1.5]$ m/s in a period of time that is uniformly distributed between $[20, 40]$ minutes. Participants move over a street network in a random walk fashion between a random generated starting and arrival point respecting the walking time constraint. The number of users is set to 10000, unless otherwise stated. The current battery charge is generated considering a value of full capacity and an initial level. The full battery capacity is variable following most popular models of smartphones available on the market and is randomly picked from a list including 2200 mAh (Huawei P8 Lite), 2550 mAh (Samsung Galaxy S6), 2800 mAh (LG G5) and 3300 mAh (Samsung Galaxy J7). The initial battery level when users start walking is uniformly distributed in the range $[10 - 90]\%$.

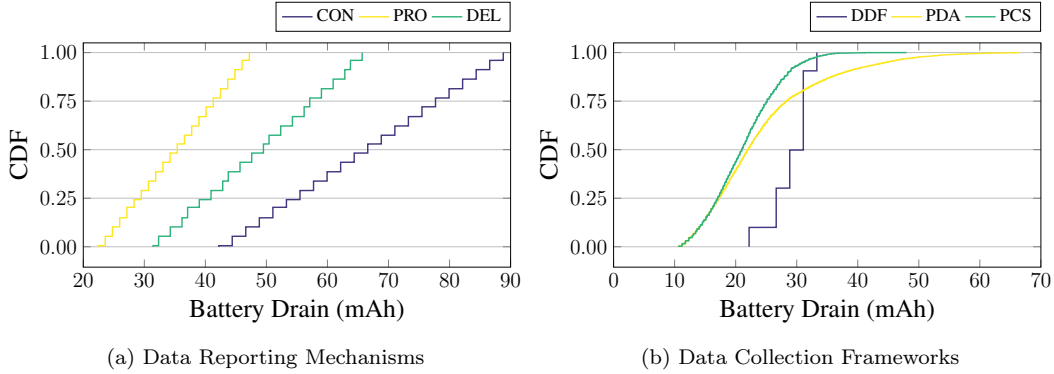


Figure 9: CDF average battery drain per user on large scale

5.2. Simulation Results

For performance evaluation, we conduct experiments to assess the energy consumption and the amount of collected data in different urban environments and the fairness of each DCF.

Energy Consumption. Fig. 9 presents the CDF of the per-user battery drain for the proposed DRMs and the DCFs under analysis in Luxembourg City. Interestingly, the difference between DRMs and DCFs is substantial. The reason is that DRMs lack of important components of DCFs such as a feedback from collector on data utility or a criteria to stop contribution. Fig. 9(a) shows the CDF of battery drain for the DRMs, which differ by range of values and slope. The steps in the profiles represent groups of users that have achieved a certain amount of battery drain. As expected, CON is the most energy consuming DRM and exhibits the highest range variability while PRO and not DEL is the less consuming DRM. However, implementing expensive DRM like CON in a DCF that can be tuned to limit user contribution like DDF is beneficial (see Fig. 9(b)). With DDF, all the mobile devices spend 33 mAh at maximum. In comparison, the percentage of users that spend more than 33 mAh for with PCS and PDA is respectively 3% and 16%. DDF reduces the percentage of mobile devices with low battery drain. This fact highlights that the organizer of a campaign employs effectively participants that accepted to join the campaign for gathering information and are rewarded on this basis. On the other hand, a substantial number of devices consume a little amount of battery with PCS. With PDA, most users consume a bigger amount of energy compared to PCS, but it is due to an increase of the amount of contributed data, as will be later highlighted in Fig. 13 and in Fig. 14.

Fig. 10 illustrates the similarity of battery drain across different urban environments, showing that the urban morphology of street networks and the size of the city have a minor impact on the energy performance of the DCFs. DDF exhibits a CDF that mimics a step function, where each step indicates a group of participants that has contributed to the collector a similar amount of data. Hence, they have stopped to acquire data for one of the possible motivation: (i) the battery drain due to sensing

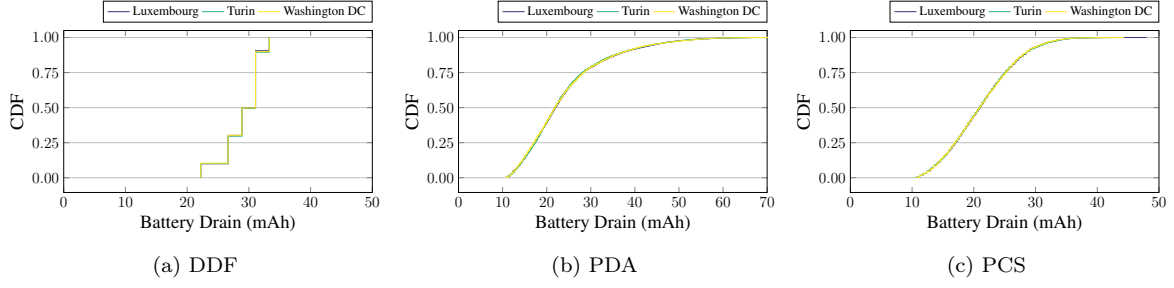


Figure 10: CDF of battery drain per user for considered DCFs in different cities

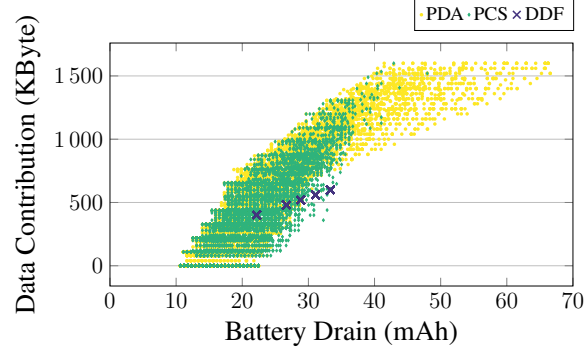


Figure 11: Amount of collected data and the associated battery drain in Luxembourg City

and reporting operations has exceeded a given value, (ii) users already reported a fixed amount of data.

460 Fig. 11 presents the amount of reported information and the associated energy consumption for each DCF. Marks in the graphic represent the battery drain that a group of participants has consumed to contribute a certain amount of data. Interestingly, DDF presents a low number of marks. The motivation is that mobile devices show a similar behavior because of the stopping mechanism which indirectly controls the energy consumption. On the other side, PDA and PCS show much higher variability due to the different reporting mechanisms: to contribute a certain amount of data, the participants spend
465 different amounts of energy. This variability increases when the total contribution becomes higher. From a practical point of view, simulations exhibit that applying an incentive mechanism based only on the amount of acquired information misses to reward fairly individuals because of the technical implementation of reporting policies.

470 **Amount of Collected Data.** The next experiments assess the amount of contributed data. Fig. 12 illustrates the trajectories of 5 participants walking in Luxembourg City. The aim is to focus on the periods of active contribution and show the differences between the DCFs. DDF allows the users to perform continuous contribution until the mobile devices of the participants meet the conditions to stop to contribute data. PDA shows an intermittent reporting, which is attributed to the probability
475 of performing data transmission. For example, the sections of the trajectories when users perform

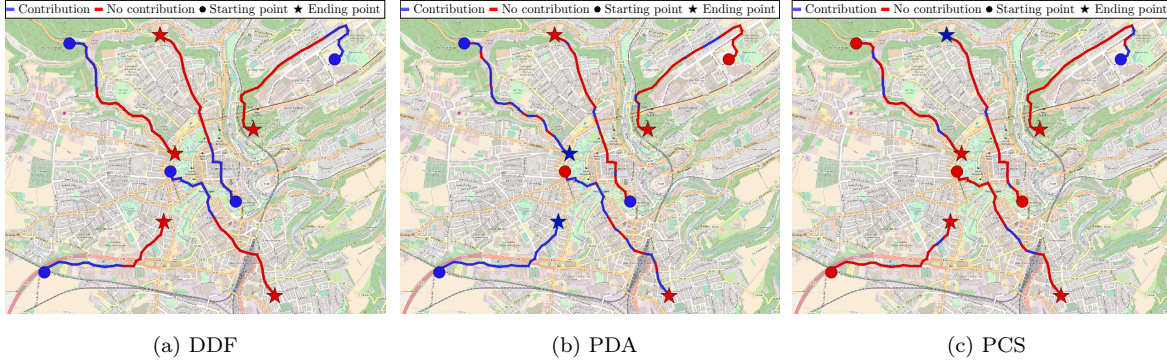


Figure 12: User trajectories with the associated data contribution in Luxembourg City

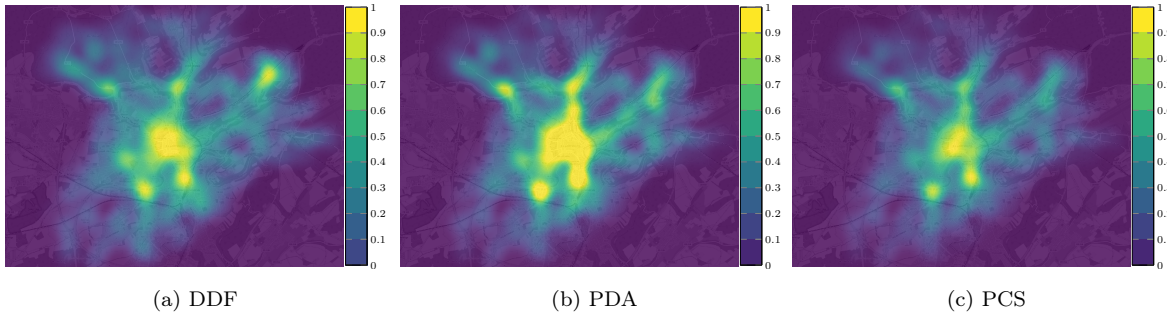


Figure 13: Heatmaps of Luxembourg City with different DCFs

continuously active contribution depend on the feedback of the collector that asks for additional data. Hence, for all the users, the probability to transmit is high. With PCS, the user active contribution is extremely limited and totally depends on the probability of performing user calls and their duration.

Fig. 13 shows the spatial distribution of the total amount of collected data at the end of the simulation period for Luxembourg City. The heatmap is normalized between 0 and 1 and 1 indicates a total of 100 MB of data generated during the entire simulation period. PDA achieves a high spatial distribution of amount of collected data and this is because it collects data until the collector has gathered a sufficient amount of data and lowers the transmission probability of the users. DDF shows a lower amount of collected data in the center due to the stopping mechanism as shown in Fig. 9b. PCS achieves the lowest amount of contributed data. Indeed, although users perform continuous sensing, data reporting fully depends on the probability of performing phone calls.

Fig. 14 shows the amount of contribution in Luxembourg City comparing the considered DCFs for different number of users. DDF is able to gather a considerable amount of data due to continuous reporting even if users stop to contribute to save energy. PDA achieves a similar amount of contributed data of DDF. On the contrary, PCS achieves the lowest amount of acquired information and would most likely fail to accurately monitor area of interests. Again, the motivation lies in the reporting mechanism implemented.

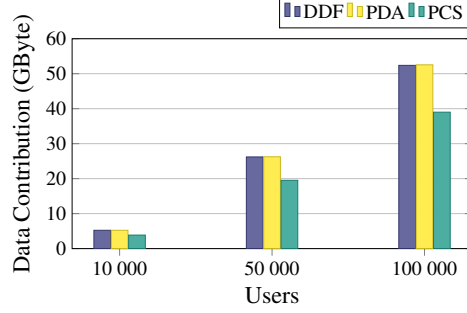


Figure 14: Amount of contributed data for considered DCFs in Luxembourg City

Fairness. Fig. 15 compares DCFs according to the different fairness indexes presented in Section 4.2. The results are obtained after 100 rounds of simulations. Fig. 15(a) shows the data contribution fairness index (F_D) in form of boxplots. DDF achieves the highest values of fairness in data contribution and the reason is twofold. First, feedback from collector regulates the amount of data generated. Second, the contribution for each participant is fair in proportion to the amount of time they walk, i.e., users that walk for longer time periods contribute more data. PCS, that depends on the probability of users phone calls is less fair than DDF. In other words, users do not contribute the same amount of data when walk for the same amount of time. PDA presents the lowest value of data contribution fairness because some of the users contribute significant amounts of data when the collector misses data, but if the urgency for additional data is low the users do not contribute at all. Note that for design choice, PDA implements the AO-S algorithm, which is a basic setting and is not as fair as the AO-F variant [31]. In AO-S the transmission probability is determined according to a satisfactory index determined by the collector, while AO-F takes into account user fairness by determining the transmission probability to balance users load. To illustrate, the transmission probability of users that have backed-off for longer periods is higher than the one of users that reported data in previous timeslots. Fig. 15(b) shows the results on the battery fairness index (F_B). If compared to the data fairness index, DDF presents much lower values because of the mechanism that stops users if they have spent a certain amount of energy for data contribution. However, this does not differentiate between users that have high or low initial battery level. For example, upon setting the stopping threshold to 1% of the battery, the impact is different if the initial level is 10% than if it was 90%. PCS and PDA are very similar and lower than DDF. Although none implements a stopping criteria, the implicit stopping method is given by the amount of time the users walk. Finally, Fig. 15(c) shows the Crowdsensing Index (F_{CS}) which combines the previous indexes together. PCS is the most uniform is due to the lack of feedbacks from the collector and sporadic contribution of each mobile device during phone calls. PDA shows a linear decrease between data and battery fairness while DDF is the opposite and exhibits a linear increase. Such behavior is due to the underlying properties of the DCFs.

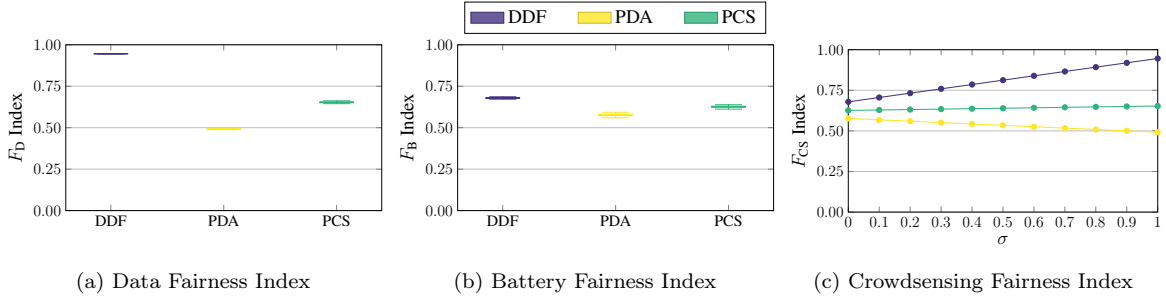


Figure 15: Fairness Indexes

6. Discussion

The design of the fifth generation (5G) of cellular mobile networks aims at supporting higher mobility, massive connectivity, higher data rates and lower latency. 5G networks will dramatically change the landscape of mobile networks and will play a significant role also in crowdsensing.

As future work, we indeed plan to investigate efficient solutions for crowdsensing applications in 5G mobile networks. The design of the fifth generation (5G) of cellular mobile networks aims at simultaneously support various type of services like enhanced mobile broadband (eMBB) and massive machine type communications (mMTC), and ultra-reliable low-latency communications (URLLC). Previous studies have focused on defining a flexible frame structure with variable duration of the basic time unit for data transmission. In 4G systems, the Transmission Time Interval (TTI) was fixed to 1 ms. While this figure can suffice for eMBB services, to accommodate URLLC services shorter TTIs are required [47]. Indeed, the fixed and long TTI duration is highly inefficient for sporadic transmission of small packets that is common in crowdsensing applications.

7. Conclusion

DCFes define the efficiency of MCS systems in terms of quality of information acquired. Effective frameworks make minimal the energy costs associated to sensing and reporting. Profiling DCFes' energy is crucial to assess the costs of a sensing campaign and to properly plan user incentives plans like monetary rewarding. This paper has made the first experimental assessment and comparison of energy consumption of DCFes. An Android application was developed for the purpose and profiled through a power monitor and Wireshark for network-related insights. The acquired energy measurements have been utilized for large-scale analysis with CrowdSenSim simulator to test effectiveness of the frameworks in Luxembourg City, Turin and Washington DC. The results show that DCFes energy consumption differ because of the data reporting mechanism implemented: DCFes with probabilistic reporting comparatively achieve higher energy consumption. Furthermore, such DCFes present high

variability, i.e., to produce the same amount of data, the associated energy cost of different users can be significantly different. Consequently, DCFs with continuous reporting that implement mechanisms to block sensing and data delivery after a certain amount of contribution are more energy-effective in harvesting data from the crowd.

Acknowledgment

The authors would like to thank Piergiorgio Vitello and the anonymous reviewers for the precious and valuable feedback on the draft.

References

- [1] N. B. Grimm, S. H. Faeth, N. E. Golubiewski, C. L. Redman, J. Wu, X. Bai, J. M. Briggs, Global change and the ecology of cities, *science* 319 (5864) (2008) 756–760. doi:10.1126/science.1150195.
- [2] D. Puiu, P. Barnaghi, R. Tönjes, D. Kümper, M. I. Ali, A. Mileo, J. X. Parreira, M. Fischer, S. Koloza, N. Farajidavar, F. Gao, T. Iggena, T. L. Pham, C. S. Nechifor, D. Puschmann, J. Fernandes, CityPulse: Large scale data analytics framework for smart cities, *IEEE Access* 4 (2016) 1086–1108. doi:10.1109/ACCESS.2016.2541999.
- [3] A. Zanella, N. Bui, A. Castellani, L. Vangelista, M. Zorzi, Internet of Things for smart cities, *IEEE Internet of Things Journal* 1 (1) (2014) 22–32. doi:10.1109/JIOT.2014.2306328.
- [4] G. Cacciatore, C. Fiandrino, D. Kliazovich, F. Granelli, P. Bouvry, Cost analysis of smart lighting solutions for smart cities, in: *IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6. doi:10.1109/ICC.2017.7996886.
- [5] L. Duan, L. Huang, C. Langbort, A. Pozdnukhov, J. Walrand, L. Zhang, Human-in-the-loop mobile networks: A survey of recent advancements, *IEEE Journal on Selected Areas in Communications* 35 (4) (2017) 813–831. doi:10.1109/JSAC.2017.2695738.
- [6] R. Ganti, F. Ye, H. Lei, Mobile crowdsensing: current state and future challenges, *IEEE Communications Magazine* 49 (11) (2011) 32–39. doi:10.1109/MCOM.2011.6069707.
- [7] G. Cardone, A. Cirri, A. Corradi, L. Foschini, R. Ianniello, R. Montanari, Crowdsensing in urban areas for city-scale mass gathering management: Geofencing and activity recognition, *IEEE Sensors Journal* 14 (12) (2014) 4185–4195. doi:10.1109/JSEN.2014.2344023.
- [8] W. Khan, Y. Xiang, M. Aalsalem, Q. Arshad, Mobile phone sensing systems: A survey, *IEEE Communications Surveys Tutorials* 15 (1) (2013) 402–427. doi:10.1109/SURV.2012.031412.00077.

- [9] Q. Wang, B. Guo, L. Wang, T. Xin, H. Du, H. Chen, Z. Yu, Crowdwatoh: Dynamic sidewalk obstacle detection using mobile crowd sensing, *IEEE Internet of Things Journal* 4 (6) (2017) 2159–2171. doi:10.1109/JIOT.2017.2750324.
- 575 [10] A. Fox, B. V. Kumar, J. Chen, F. Bai, Multi-lane pothole detection from crowdsourced under-sampled vehicle sensor data, *IEEE Transactions on Mobile Computing* 16 (12) (2017) 3417–3430. doi:doi.ieeecomputersociety.org/10.1109/TMC.2017.2690995.
- [11] S. Kim, C. Robson, T. Zimmerman, J. Pierce, E. M. Haber, Creek Watch: Pairing usefulness and usability for successful citizen science, in: *Proc. of the SIGCHI Conference on Human Factors in Computing Systems, CHI, ACM, 2011*, pp. 2125–2134. doi:10.1145/1978942.1979251.
- 580 [12] S. Reddy, V. Samanta, Urban Sensing: Garbage Watch, UCLA Center for Embedded Networked Sensing (2011).
- [13] V. Sivaraman, J. Carrapetta, K. Hu, B. G. Luxan, Hazewatch: A participatory sensor system for monitoring air pollution in sydney, in: *IEEE Conference on Local Computer Networks (LCN) - Workshops, 2013*, pp. 56–64. doi:10.1109/LCNW.2013.6758498.
- 585 [14] A. Capponi, C. Fiandrino, C. Franck, U. Sorger, D. Kliazovich, P. Bouvry, Assessing performance of internet of things-based mobile crowdsensing systems for sensing as a service applications in smart cities, in: *IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*, 2016, pp. 456–459. doi:10.1109/CloudCom.2016.0077.
- 590 [15] A. Capponi, C. Fiandrino, D. Kliazovich, P. Bouvry, Energy efficient data collection in opportunistic mobile crowdsensing architectures for smart cities, in: *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2017, pp. 307–312. doi:10.1109/INFCOMW.2017.8116394.
- [16] C. Fiandrino, F. Anjomshoa, B. Kantarci, D. Kliazovich, P. Bouvry, J. N. Matthews, Sociability-driven framework for data acquisition in mobile crowdsensing over fog computing platforms for smart cities, *IEEE Transactions on Sustainable Computing* 2 (4) (2017) 345–358. doi:10.1109/TSUSC.2017.2702060.
- 595 [17] B. Song, H. Shah-Mansouri, V. W. S. Wong, Quality of sensing aware budget feasible mechanism for mobile crowdsensing, *IEEE Transactions on Wireless Communications* 16 (6) (2017) 3619–3631. doi:10.1109/TWC.2017.2686085.
- 600 [18] F. Restuccia, N. Ghosh, S. Bhattacharjee, S. K. Das, T. Melodia, Quality of information in mobile crowdsensing: Survey and research challenges, *CoRR* abs/1709.03583.

- [19] M. Tomasoni, A. Capponi, C. Fiandrino, D. Kliazovich, F. Granelli, P. Bouvry, Profiling energy efficiency of mobile crowdsensing data collection frameworks for smart city applications, in: IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud), 2018, pp. 1–8. doi:10.1109/MobileCloud.2018.00009.
- [20] C. Fiandrino, A. Capponi, G. Cacciatore, D. Kliazovich, U. Sorger, P. Bouvry, B. Kantarci, F. Granelli, S. Giordano, CrowdSenSim: a simulation platform for mobile crowdsensing in realistic urban environments, IEEE Access 5 (2017) 3490–3503. doi:10.1109/ACCESS.2017.2671678.
- [21] Z. Zheng, Y. Peng, F. Wu, S. Tang, G. Chen, Trading data in the crowd: Profit-driven data acquisition for mobile crowdsensing, IEEE Journal on Selected Areas in Communications 35 (2) (2017) 486–501. doi:10.1109/JSAC.2017.2659258.
- [22] W. Wu, J. Wang, M. Li, K. Liu, F. Shan, J. Luo, Energy-efficient transmission with data sharing in participatory sensing systems, IEEE Journal on Selected Areas in Communications 34 (12) (2016) 4048–4062. doi:10.1109/JSAC.2016.2621359.
- [23] J. Wang, J. Tang, G. Xue, D. Yang, Towards energy-efficient task scheduling on smart-phones in mobile crowd sensing systems, Computer Networks 115 (2017) 100 – 109. doi:https://doi.org/10.1016/j.comnet.2016.11.020.
- [24] Q. Zhao, Y. Zhu, H. Zhu, J. Cao, G. Xue, B. Li, Fair energy-efficient sensing task allocation in participatory sensing with smartphones, in: IEEE INFOCOM 2014 - IEEE Conference on Computer Communications, 2014, pp. 1366–1374. doi:10.1109/INFOCOM.2014.6848070.
- [25] C. Liu, B. Zhang, X. Su, J. Ma, W. Wang, K. Leung, Energy-aware participant selection for smartphone-enabled mobile crowd sensing, IEEE Systems Journal PP (99) (2015) 1–12. doi:10.1109/JSYST.2015.2430362.
- [26] P. P. Jayaraman, J. B. Gomes, H. L. Nguyen, Z. S. Abdallah, S. Krishnaswamy, A. Zaslavsky, Scalable energy-efficient distributed data analytics for crowdsensing applications in mobile environments, IEEE Transactions on Computational Social Systems 2 (3) (2015) 109–123.
- [27] W. Sherchan, P. P. Jayaraman, S. Krishnaswamy, A. Zaslavsky, S. Loke, A. Sinha, Using on-the-move mining for mobile crowdsensing, in: IEEE 13th International Conference on Mobile Data Management, 2012, pp. 115–124.
- [28] L. Wang, D. Zhang, Z. Yan, H. Xiong, B. Xie, effSense: A novel mobile crowd-sensing framework for energy-efficient and cost-effective data uploading, IEEE Transactions on Systems, Man, and Cybernetics: Systems 45 (12) (2015) 1549–1563. doi:10.1109/TSMC.2015.2418283.

- [29] J. Peng, Y. Zhu, Q. Zhao, H. Zhu, J. Cao, G. Xue, B. Li, Fair energy-efficient sensing task allocation in participatory sensing with smartphones, *The Computer Journal* 60 (6) (2017) 850–865.
- [30] A. Capponi, C. Fiandrino, D. Kliazovich, P. Bouvry, S. Giordano, A cost-effective distributed framework for data collection in cloud-based mobile crowd sensing architectures, *IEEE Transactions on Sustainable Computing* 2 (1) (2017) 3–16. doi:10.1109/TSUSC.2017.2666043.
- [31] F. Montori, L. Bedogni, L. Bononi, Distributed data collection control in opportunistic mobile crowdsensing, in: *Proc. of the 3rd Workshop on Experiences with the Design and Implementation of Smart Objects, SMARTOBJECTS*, ACM, 2017, pp. 19–24. doi:10.1145/3127502.3127509.
- [32] N. D. Lane, Y. Chon, L. Zhou, Y. Zhang, F. Li, D. Kim, G. Ding, F. Zhao, H. Cha, Piggyback crowdsensing (PCS): Energy efficient crowdsourcing of mobile sensor data by exploiting smart-phone app opportunities, in: *11th ACM Conference on Embedded Networked Sensor Systems, SenSys*, 2013, pp. 1–14. doi:10.1145/2517351.2517372.
- [33] N. Haderer, R. Rouvoy, L. Seinturier, A preliminary investigation of user incentives to leverage crowdsensing activities, in: *IEEE International Conference on Pervasive Computing and Communications Workshops*, 2013, pp. 199–204. doi:10.1109/PerComW.2013.6529481.
- [34] F. Paganelli, S. Turchi, D. Giuli, A web of things framework for RESTful applications and its experimentation in a smart city, *IEEE Systems Journal* 10 (4) (2016) 1412–1423. doi:10.1109/JSYST.2014.2354835.
- [35] R. Mulero, A. Almeida, G. Azkune, L. Mainetti, V. Mighali, L. Patrono, P. Rametta, I. Sergi, An AAL system based on IoT technologies and linked open data for elderly monitoring in smart cities, in: *2nd International Multidisciplinary Conference on Computer and Energy Science (SpliTech)*, 2017, pp. 1–6.
- [36] F. Kalim, J. P. Jeong, M. U. Ilyas, Crater: A crowd sensing application to estimate road conditions, *IEEE Access* 4 (2016) 8317–8326. doi:10.1109/ACCESS.2016.2607719.
- [37] J. W. Sung, S. J. Han, Data bundling for energy efficient communication of wearable devices, in: *IEEE 40th Conference on Local Computer Networks (LCN)*, 2015, pp. 321–328. doi:10.1109/LCN.2015.7366326.
- [38] R. LiKamWa, Z. Wang, A. Carroll, F. X. Lin, L. Zhong, Draining our glass: An energy and heat characterization of google glass, in: *Proc. of 5th Asia-Pacific Workshop on Systems, APSys*, ACM, 2014, pp. 1–7. doi:10.1145/2637166.2637230.

- [39] J. C. V. Bedregal, E. G. C. Gutierrez, R. E. A. M., Optimizing energy consumption per application in mobile devices, in: International Conference on Information Society, 2013, pp. 106–110.
- [40] G. P. Perrucci, F. H. P. Fitzek, J. Widmer, Survey on energy consumption entities on the smart-phone platform, in: IEEE 73rd Vehicular Technology Conference (VTC Spring), 2011, pp. 1–6. doi:10.1109/VETECS.2011.5956528.
- [41] P. Serrano, A. Garcia-Saavedra, G. Bianchi, A. Banchs, A. Azcorra, Per-frame energy consumption in 802.11 devices and its implication on modeling and design, IEEE/ACM Transactions on Networking 23 (4) (2015) 1243–1256.
- [42] K. Tan, J. Fang, Y. Zhang, S. Chen, L. Shi, J. Zhang, Y. Zhang, Fine-grained channel access in wireless lan, in: Proc. of the ACM SIGCOMM, ACM, 2010, pp. 147–158. doi:10.1145/1851182.1851202.
- [43] S. Chessa, M. Girolami, L. Foschini, R. Ianniello, A. Corradi, P. Bellavista, Mobile crowd sensing management with the ParticipAct living lab, Pervasive and Mobile Computing 38 (2017) 200 – 214. doi:https://doi.org/10.1016/j.pmcj.2016.09.005.
- [44] R. Jain, D. Chiu, W. Hawe, A quantitative measure of fairness and discrimination for resource allocation in shared computer systems, CoRR cs.NI/9809099.
- [45] P. Vitello, A. Capponi, C. Fiandrino, P. Giaccone, D. Kliazovich, P. Bouvry, High-precision design of pedestrian mobility for smart city simulators, in: IEEE International Conference on Communications (ICC), 2018, pp. 1–6. doi:10.1109/ICC.2018.8422599.
- [46] G. Boeing, OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks, Computers, Environment and Urban Systems 65 (2017) 126–139. doi:https://doi.org/10.1016/j.compenvurbsys.2017.05.004.
- [47] K. I. Pedersen, G. Berardinelli, F. Frederiksen, P. Mogensen, A. Szufarska, A flexible 5G frame structure design for frequency-division duplex cases, IEEE Communications Magazine 54 (3) (2016) 53–59. doi:10.1109/MCOM.2016.7432148.

Biographies



Mattia Tomasoni is currently a MS student in Telecommunications Engineering at the University of Trento. He is performing his master thesis as an intern at the University of Luxembourg.



Andrea Capponi is a Ph.D. candidate at the University of Luxembourg. He received the Bachelor Degree in Telecommunication Engineering and the Master Degree in Telecommunication Engineering both from the University of Pisa, Italy. His primary research interests are in the field of mobile crowdsensing, Internet of Things (IoT), and smart cities.



Claudio Fiandrino joined as a postdoctoral researcher the IMDEA Networks Institute in December 2016 right after having obtained his Ph.D. degree at the University of Luxembourg. He received the Bachelor Degree in Ingegneria Telematica in 2010 and the Master Degree in Computer and Communication Networks Engineering in 2012 both from Politecnico di Torino. Claudio also holds the 2016 SmartICT Certificate on standardization for business innovation from the joint program of University of Luxembourg and ILNAS, the National Standardization Agency. Claudio's work on indoor localization over fog computing platforms received the Best Paper Award in IEEE CloudNet 2016. He is member of IEEE and ACM, and served as Publication and Web Chair at IEEE CloudNet 2014 and as Publicity Chair in ACM/IEEE ANCS 2018. His primary research interests include mobile crowdsensing, multi-access edge/fog computing and ultra-reliable and low latency communications.



Dzmitry Kliazovich is a Head of Innovation at ExaMotive. He was a Senior Scientist at the Faculty of Science, Technology, and Communication of the University of Luxembourg. Dr. Kliazovich holds an award-winning Ph.D. in Information and Telecommunication Technologies from the University of Trento (Italy). His works on cloud computing, energy-efficiency, indoor localization, and mobile networks received IEEE/ACM Best Paper Awards. He coordinated organization and chaired a number of highly ranked international conferences and symposia, including the IEEE International Conference on Cloud Networking (CloudNet 2014). He is the Associate Editor of the IEEE Communications Surveys and Tutorials and of the IEEE Transactions of Cloud Computing journals. His main research activities are in the field of intelligent transportation systems, telecommunications, cloud computing, and Internet of Things (IoT).



Fabrizio Granelli is Associate Professor and Delegate for Education at the Dept. of Information Engineering and Computer Science (DISI) of the University of Trento (Italy) and IEEE ComSoc Director for Online Content. He received the «Laurea» (M.Sc.) degree in Electronic Engineering from the University of Genoa, Italy, in 1997, with a thesis on video coding, awarded with the TELECOM Italy prize, and the Ph.D. in Telecommunications from the same university, in 2001. Since 2000 he is carrying on his research and didactical activities (currently Associate Professor in Telecommunications) at the Dept.

of Information Engineering and Computer Science – University of Trento (Italy) and coordinator of the Networking Laboratory. Between 2004 and 2015 for a total of six months, he was visiting professor at the State University of Campinas (Brasil). In 2016, he was visiting professor at the University of Tokyo (Japan). He is Associate Editor-in-Chief of IEEE Communications Surveys and Tutorials journal and he was IEEE ComSoc Distinguished Lecturer for the period 2012-15 (2 terms). He is author or co-author of more than 170 papers published in international journals, books and conferences, focused on networking.



Pascal Bouvry is a professor in the Computer Science and Communication research unit of the Faculty of Science, Technology and Communication at the University of Luxembourg and a faculty member at the Luxembourg Interdisciplinary Center of Security, Reliability, and Trust. His research interests include cloud & parallel computing, optimization, security and reliability. Prof. Bouvry has a Ph.D. in computer science from the University of Grenoble (INPG), France. He is on the IEEE Cloud Computing and Elsevier Swarm and Evolutionary Computation editorial boards. He is also acting as communication vice-chair of the IEEE STC on Sustainable Computing and co-founder of the IEEE TC on Cybernetics for Cyber-Physical Systems. A full biography is available on page <http://pascal.bouvry.org>. Contact him at pascal.bouvry@uni.lu.