Mobile Crowd Sensing Management with the ParticipAct Living Lab

Stefano Chessa[a], Michele Girolami[a,b], Luca Foschini[c], Raffaele Ianniello[c], Antonio Corradi[c], Paolo Bellavista[c,*]

[a] *Department of Computer Science and University of Pisa, Pisa, Italy*
[b] *ISTI-CNR, Pisa, Italy*
[c] *Dipartimento di Informatica: Scienza e Ingegneria Università di Bologna, Bologna, Italy*

Abstract

The widespread availability of smartphones have enabled the blossom of Mobile Crowd Sensing (MCS) projects, whose goal is to involve users in participatory tasks aimed at building large real-world datasets. In this framework, we present the large-scale experience of the ParticipAct Living Lab, an ongoing experiment at the University of Bologna, which involves about 170 students in MCS campaigns. Specifically, we originally present the analysis of the large set of ParticipAct collected results against some primary datasets in the literature; we present the evaluation and assessment of the original participatory sensing campaign management aspects of ParticipAct; and we report the lessons learned from this wide-scale deployment experience.

Keywords: Mobile Systems, Mobile Social Networks, Mobile Crowd Sensing Management

2010 MSC: 00-01, 99-00

---

*Corresponding author

Email addresses: stefano.chessa@di.unipi.it (Stefano Chessa), michele.girolami@isti.cnr.it (Michele Girolami), luca.foschini@unibo.it (Luca Foschini), raffaele.ianniello@unibo.it (Raffaele Ianniello), antonio.corradi@unibo.it (Antonio Corradi), paolo.bellavista@unibo.it (Paolo Bellavista)

1. Introduction

The large availability of mobile devices with sensing capabilities, combined with the pervasive availability of communication infrastructures, gave rise, in recent years, to a number of platforms for Mobile Crowd Sensing (MCS). MCS is commonly referred to as a paradigm for distributed gathering of heterogeneous sensing data from pocket devices used by the crowds. Different recent projects produced various significant datasets by using MCS platforms [1, 2, 3, 4, 5, 6, 7, 8], obtained in different geographical areas, with different objectives and based on different sensing technologies. Among these, it is worth mentioning in particular the Cambridge [1] and the MIT reality [2] campaigns (conducted in 2005/06), the Mobile Data Challenge Nokia (MDC Nokia) [4, 5] campaign conducted with more powerful smartphones in 2009 and, more recently, the ParticipAct one [3] that is currently running since 2013.

These datasets are highly valued for the research community, as they provide a way to test, assess, and compare differentiated solutions for diverse application scenarios ranging from MCS [9, 10] to mobile social computing [6, 11] and opportunistic networking [12], based not only on real-world traces of human mobility but also on evidences of their activities and social behaviors. In particular, we choose datasets that provide information about the user mobility (either in the form of co-location traces or GPS coordinates) and datasets providing information about the data gathered by user's devices.

However, most research effort so far has been devoted to dataset production (including the efficient design of MCS platforms and the optimized conduction of data collection campaigns), while a relatively minor effort has been devoted to the assessment of the datasets themselves. Specifically, existing datasets address different scenarios in terms of time span of the experimentation, extension of the area of data collection, classes of involved users, capabilities of the hardware/software devices and sensing platforms in use, etc. They involve a variable number of users recruited with different modalities and with different ultimate goals for the different sensing campaigns. The objective of this work is to exploit the availability of the novel ParticipAct dataset, to make an original further step towards the assessment and quantitative comparison of primary datasets in the literature and of interest for the ParticipAct experience.

In particular, the contribution of this work is twofold. We first present a novel comparative analysis of the Cambridge, MIT Reality, MDC Nokia, and ParticipAct datasets with respect to some mobility metrics given along with the temporal dimension whose goal is to assess the capability of the datasets to reflect the human dynamics in real-world MCS scenarios. Second, we focus on the ParticipAct dataset to assess its performance as crowdsensing platform along three dimensions: user assignment policies, task acceptance (i.e., how sensing tasks are actually accepted and taken into consideration by assigned users), and task completion rate. Note that our intention is not to rank datasets, rather, it is to report lessons learned from the ParticipAct experience and from its comparison against some popular and widely accepted datasets in the related literature. We believe that this effort could be valuable for the community of researches and practitioners on the filed, by providing them with useful guidelines for the planning and implementation of future participatory sensing campaigns, also not ParticipAct-based.

The reminder of the paper is structured as follows. In Section 2 we provide a quick overview of the ParticipAct platform, in Section 3 we compare the most relevant MCS datasets available in the literature, in Section 4 we present a deep analysis of the ParticipAct dataset, Section 5 presents an overall discussions of the datasets analyzed and Section 6 concludes the paper.

2. ParticipAct at a glance

Here, for the sake of self-containment and full understandability of the remainder of the paper, we provide a very rapid overview of a few central elements of the ParticipAct MCS platform; interested readers could refer to [3] for additional details about the efficient design and implementation of ParticipAct.

MCS platforms typically adopt a client-server architecture including a client,

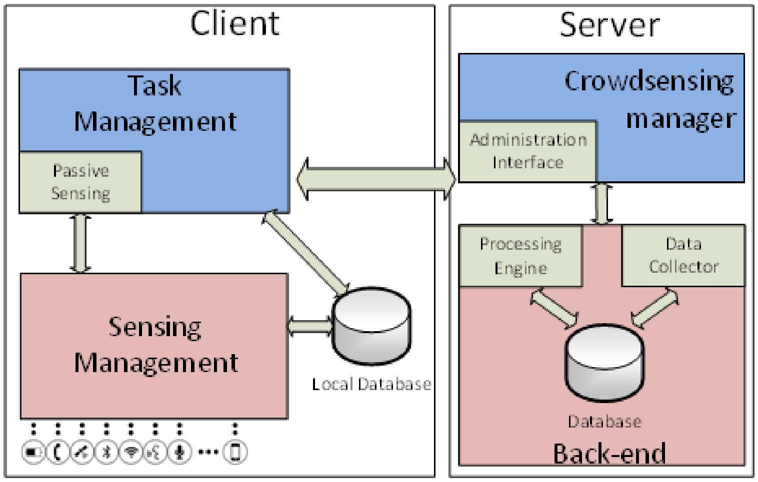


Figure 1: ParticipAct architecture.

running on user devices to manage tasks and to run all required sensing activities interacting with participants via their smartphones, and a server to store and present collected results [9].

The ParticipAct client (see Figure 1) is the component that takes care of receiving tasks, asking users whether they want to run them, managing data collection, and uploading results. Functionally, the ParticipAct client consists of two main components: the task management component and the sensing management component. These components are responsible for both interacting with users and accessing smartphone sensors. In particular, the task management component takes care of overseeing the whole MCS task life-cycle on smartphone, from managing tasks to provide users with an interface to control task execution, including the possibility to stop any sensing activity to preserve user privacy, to the final upload of sensed data. Through the sensing management component, instead, it is possible to efficiently access all the sensors available on smartphones and collecting/processing their output.

The ParticipAct server provides advanced management, storage, and analysis features for data gathered during a crowdsensing campaign. At the highest level, it comprises two main parts (see Figure 1): the back-end and the crowdsensing manager. The back-end takes care of receiving, storing, and processing sensed data, while the crowdsensing manager provides the administration interface to design, assign, and deploy sensing tasks. In a more detailed view, the back-end realizes the needed communication functions to exchange tasks and receive results, and manages the whole data life-cycle. The crowdsensing manager, instead, is the administrator-facing part of ParticipAct that exploits back-end exported functions to provide easy-to-use and highly configurable administration features. The Web administration interface allows full administration of the whole crowdsensing, including management of user profiles, task design and the definition of their assignment strategies, and data review.

Focusing on the MCS model, ParticipAct supports two main types of sensing tasks. Passive sensing tasks enable automatic collection of smartphone generated data without user intervention, such as accelerometer, WiFi scans, and ambient noise level. Active sensing tasks, instead, contemplate active user contribution, such as taking a photo, answering a survey, and tagging a place. In addition, recognizing the importance to link MCS activities to physical places, ParticipAct enables a highly flexible model that allows defining where the task will be notified and executed. Geonotification associates tasks to one or more geographical areas and automatically notifies them to users as they enter those areas. Geoexecution, instead, associates tasks to one or more geographical areas and enables task execution, e.g., data collection, only when users are located therein.

Additional details about client and server side components of ParticipAct, out of the scope of the present paper, are available in [3, 10].

3. Analysis of Mobility Metrics

To validate the ParticipAct dataset, we assess it against a number of metrics widely used for mobility studies and a selected set of datasets available in the literature, which are representative of different mobility scenarios [13, 14, 15].

Specifically, we consider Cambridge [1], MIT Reality [2] and Mobile Data Challenge Nokia (MDC Nokia) [4, 5], which are widely used to test social interactions in mobile applications. In particular, MDC Nokia refers to scenarios very similar to those targeted by ParticipAct: it is collected in urban/rural areas over almost one year, involves a large number of heterogeneous users, and deals with crowdsensing tasks and their assignment. Cambridge and MIT Reality instead, represent quite different scenarios and examples of measurement campaigns: they are bound in space (limited to university campus) and involve a homogeneous set of users (students with similar profiles). Cambridge is also limited in time (around two weeks), while MIT reality spans for about 8 months.

As already stated, the original goal of this paper is twofold: i) to characterize the new ParticipAct dataset by comparing it with the other three datasets considered, and ii) to draw some general inferences about the usability and level of realism of those datasets, e.g. to the purpose of their usage as traces for evaluating research solutions for MCS.

3.1. Experimental Datasets

Here we briefly summarize the main features of the selected datasets, by highlighting some aspects related to the co-location of users and to their sensing capabilities, which are specifically relevant for MCS evaluation and assessment of MCS task assignment strategies.

The **Cambridge** dataset reproduces the mobility of 36 students in the Cambridge University Campus (UK) for 12 days. Each participant carried an Intel iMote device with a Bluetooth transceiver (with a transmission range of about 30m). Each device performed a sampling scan on the Bluetooth radio with a periodicity of 120s. The only information provided by this (limited) dataset are traces of co-location among users. Note that the dataset also includes some stationary devices that are not relevant to the purpose of the MCS-oriented analysis and comparison.

The **MIT Reality** dataset reproduces the mobility of 94 students in the MIT University Campus (USA) for 246 days, in the period September 2004–January

2005. The participants used a smartphone application that tracked co-location with other users (by means of the Bluetooth radio) and other data (such as sent/received SMS, phone calls, and personal data), but no sensor data. The sampling scan period on the Bluetooth radio is quite coarse-grained, i.e. 300s.

The **MDC Nokia** dataset, collected from 2009 to 2011, involved 185 users in the Lake Geneva region (CH). The users carried a Nokia N95 phone, with an application that periodically collected several data, such as GPS, Bluetooth sightings, visited places, SMS and phone calls, and other sensor data. The sampling period for all the collected data (included GPS and Bluetooth traces) was 600s.

We extract the co-location traces from MDC Nokia both from the Bluetooth sightings and from the GPS coordinates (referred below as MDC Nokia BT and MDC Nokia GPS). In Section 3.2.1 we discuss how the use of GPS or Bluetooth affects the considered evaluation metrics.

The **ParticipAct** dataset reproduces the mobility of 173 students in the Emilia Romagna region (Italy). The data collection campaign, which is still running, began in December 2013. In this work, we consider a period of 15 months, from December 2013 to February 2015. The participants use Android smartphones with an application that runs dynamically assigned MCS tasks, as rapidly overviewed in Section 2. The application tracks the location of the device by using the Google location APIs (by fusing GPS and WiFi Hot Spot coordinates for higher precision and accuracy). The sampling scan period is 150s. Differently from the other datasets, we extract the co-location of users from device position traces: we assume that two devices are co-located and able to communicate if they are placed within 10m from each other for at least 150s. This assumption introduces a certain degree of inaccuracy, since two devices that are assumed to be co-located with this method might, in fact, not communicate. Note also that the mobility in ParticipAct is unrestricted: users live in town or sub-urban areas; some of them commute daily by train, while others walk or move by bike.

3.2. Datasets Analysis

We compare the experimental datasets with respect to several metrics commonly used when analyzing human mobility in Mobile Social Networks (MSN)[13, 14, 15]. The evaluation metrics we selected are given along with the temporal dimension, as for example we compute the number of contacts, contact duration, and inter-contact time as evolving over time (functions of time). The selected metrics are relevant for evaluating people mobility and hence they can be used to assess the capability of the datasets to capture dynamic human behaviors in a real-world MCS scenarios. Note that our goal is not that of ranking the datasets according to these metrics, also because such ranking would be unfair since the datasets have been collected in different time frames, with different sensing technologies and for different purposes. Rather, our aim is to identify the most important features of the datasets and to summarize which social sensing scenarios the datasets best describe, for example to the purpose of their exploitation in the evaluation and assessment of MCS solutions that will address specific mobility/social scenarios.

In this analysis, we focus on users co-location that is a compact representation of the dynamic encounter graph obtained from a dataset, and that is indicative of both mobility and sociality of users involved in the experimental datasets. The co-location traces report the start/end time of encounters between users (i.e., between their carried devices). In the considered datasets (with the exception of ParticipAct, as discussed in Section 3.1), the co-location traces are obtained by identifying other devices in Bluetooth/WiFi radio range.

3.2.1. Evaluation Metrics

The **number of contacts** represents the distribution of users along the time and in particular with hour-grained resolution. It gives an indication of the social activity of a user, by providing information about when individuals meet and with how many people. Moreover, as discussed in [6], the number of contacts among people is a first indicator of the mobility pattern characterizing the dataset. For example, datasets with a distribution of contacts that is very high during working hours and absent after such hours only partially reflect the human dynamics. Differently, datasets with a distribution of contacts that are evenly distributed are more realistic.

The **Contact Duration** metric measures the average duration of encounters. Contact duration can be understood as a rough estimator of the familiarity degree among people. In fact, as a general rule, the more people have long-lasting contacts, the more they are involved in a non-occasional social relationship. This is the case for friends or relatives that tend to stay in contact longer with respect to occasional gatherings. Nevertheless, the analysis of the contact duration only reveals the general trend of the dataset of reproducing gathering among unknown people of among acquaintances.

The **Inter-Contact Time (ICT)** measures the time elapsed between two consecutive encounters for a pair of individuals. It is a good indicator of the frequency of encounters and, hence, of how much a dataset can capture a dynamic scenario. In datasets with very short ICT, people enter in contact very frequently with the same people; this is the case, for example, of datasets bound to indoor locations and with a restricted number of participants.

The **total number of encounters** and the **unique number of encounters** measure, for each user, the number of other users met in a period and the number of other users met only once in the same period, respectively. These metrics are indicative of the social attitude of a person in entering in contact with many other people, or conversely to interact only with a few people.

The **amount of data points** measures the number of active users of ParticipAct and MDC Nokia at very different conditions. In particular, for those datasets we identified a crowded period during which many users are supposed to contribute to the MCS campaign with respect to a more isolated period during which few users are active. The goal of such comparison is to assess the capability of such datasets of reproducing the natural rhythm of humans for a MCS campaign.

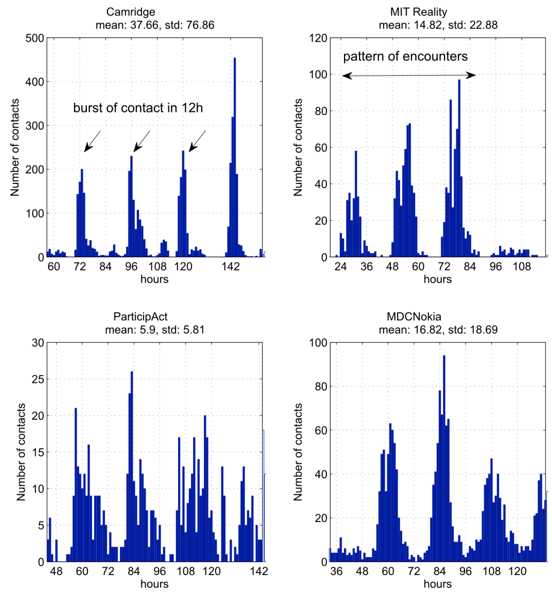


Figure 2: Number of hourly contacts.

3.2.2. Experimental Results

Results concerning the number of hourly contacts in a period of 7 days are shown in Figure 2. The histograms show that all the datasets have a similar trend in the distribution of the contacts along the day. In particular, people tend to meet other people in bursts during the daily hours. Moreover, the number of encounters repeats over the days, giving rise to an intuitive and expected pattern of encounters. In the Cambridge and the MIT Reality the bursts are denser in the first part of the daily hours (consider that these datasets are bound to university campuses where people interact during specific working hours). In the ParticipAct and the MDC Nokia datasets, instead, the bursts of contacts cover all the daily hours. Hence, a first consideration is that MDC Nokia as well as the ParticipAct datasets better capture social interactions happening also after regular working hours.

Figure 3 shows in more detail the number of contacts per hour of ParticipAct

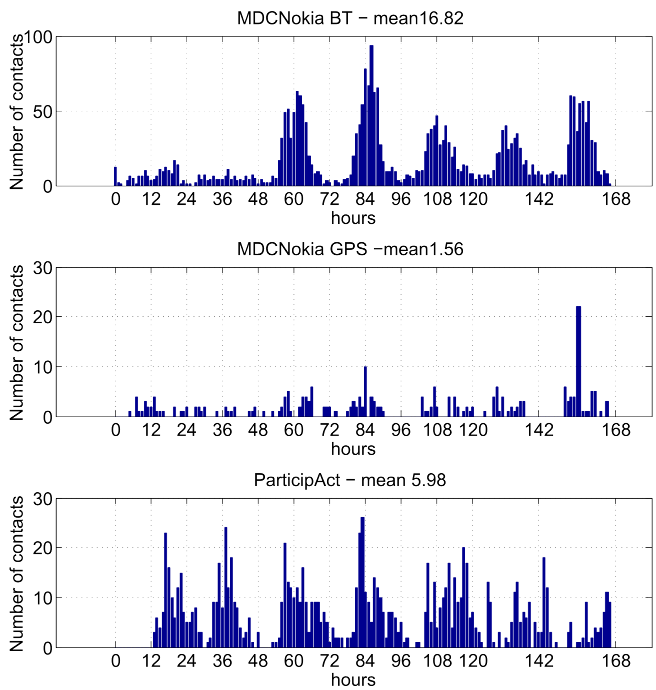


Figure 3: Average number of contacts along the time.

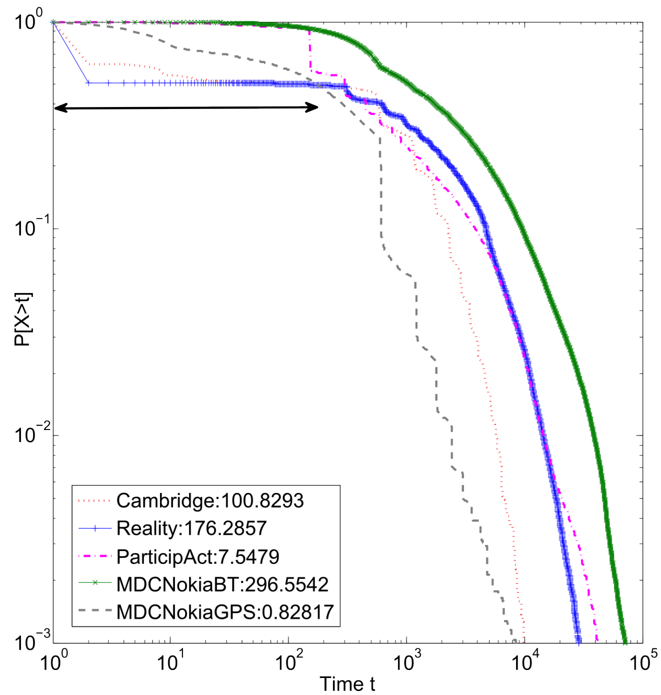


Figure 4: CCDF of contacts duration (s).

and MDC Nokia BT/GPS. The co-location traces of MDC Nokia BT better reflect the encounters among people with respect to MDC Nokia GPS. This is motivated by the fact that the traces obtained with GPS do not provide meaningful information when users are located indoor. On the other hand, the figure also shows that the limited performance given by the use of GPS can be partially overcome by benefiting from the fusion of several position systems. In particular, this occurs in the case of ParticipAct (that exploits both GPS and WiFi hotspot coordinates), which achieves an intermediate performance in terms of number of contacts between MDC Nokia BT and GPS.

Figure 4 reports the complementary cumulative distribution function (CCDF) of the duration of contacts for all the datasets in a logarithmic scale. The curves show that as the duration t increases, the probability of having contacts greater than t decreases. Such decrease is slow in the $[0-100]s$ interval, after which it follows an exponential decay rule. The dataset with the shortest contact dura-

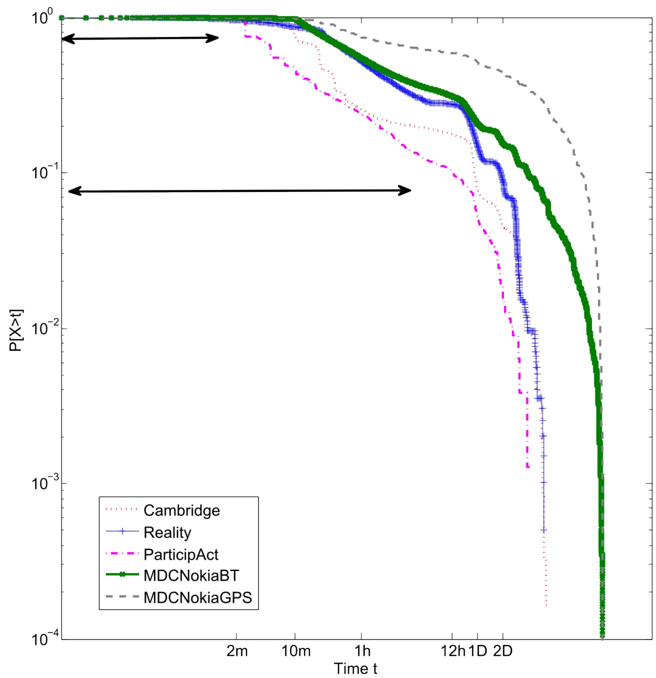


Figure 5: CCDF of inter-contact times (s).

tion is MDC Nokia GPS (with an average of 0.82 s), while MDC Nokia BT is the dataset with the highest (with an average of 296.5 s). The other datasets, namely Cambridge, MIT Reality and ParticipAct, are bound between Nokia datasets. Also in this case, the use of the GPS coordinates only for determining the co-location traces (as for MDC Nokia GPS) introduces inaccuracy. In Figure 4 this aspect is particularly evident, since MDC Nokia BT and MDC Nokia GPS are respectively the upper and lower bound of the contacts duration.

Figure 5 shows the CCDF of the inter-contact times for any pair of devices. All the datasets have a similar CCDF trend. In particular, their CCDF follows a power-low up to roughly 12 hours, after which it decays exponentially (as also observed in [7]). In all the cases, the inter-contact time is greater than 120s with high probability, but the curves assume different behaviors after 120s. In particular, ParticipAct users tend to meet the same people more frequently, while users of MDC Nokia (both Bluetooth and GPS) meet the same less often.

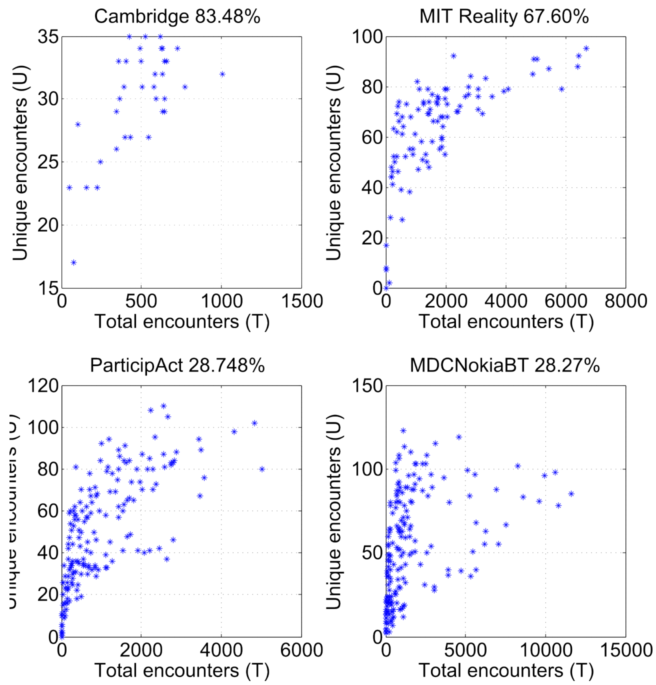


Figure 6: Heterogeneity of encounters.

Cambridge and MIT Reality are sandwiched between ParticipAct and MDC Nokia.

Figure 6 represents the heterogeneity of encounters [16], by plotting a point for each user at the coordinates given by the total number of encounters (x axis) and the number of unique encounters (y axis). From the figure, it is seen that in Cambridge and MIT Reality, which are collected in limited geographical areas, users meet many other users more often. This is confirmed by the distribution of the points. In fact, many points are placed in the upper part of the diagram that characterizes people with many unique encounters. In the case of MIT Reality, a number of points are also placed in the upper right corner; hence, in this case, people have also many total encounters. The percentage of the population visited by every person in Cambridge and MIT Reality datasets is of respectively 83.48% and 67.60%. Differently, ParticipAct and MDC Nokia BT datasets have a distribution of the points shifted in the lower left corner. This area

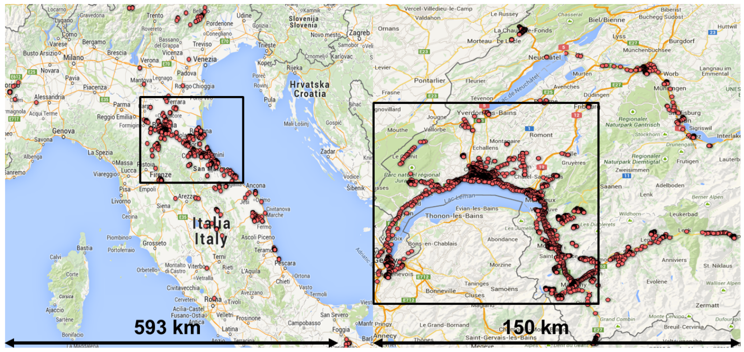


Figure 7: Geographical extension of ParticipAct and MDC Nokia datasets.

characterizes people who have few total encounters and few unique encounters. ParticipAct and MDC Nokia BT well reproduce the fact that people, in wide geographical areas, have encounters with a limited number of individuals (an average of 28.74% and 28.27% respectively).

Finally, we compare the ParticipAct and the MDC Nokia datasets along with two dimensions, namely the geographical extension and the amount of data points. Concerning the geographical extension we show in Figure 7 the user's position of ParticipAct and MDC Nokia during 1 week. As previously described, the Emilia Romagna and the Lake Geneva regions are the most crowded areas (denoted as black boxes in Figure 7), however, involved people roam also outside such regions. For example, the figure shows that ParticipAct users traveled along the entire north/center of Italy, covering over 500Km, while the maximum distance among users in MDC Nokia is limited to around 150Km.

About the amount of data points, we compare the number of active users of ParticipAct and MDC Nokia during an overloaded time period, with many users contributing to the MCS campaign, with respect to time period in which many users are not so active (i.e., summer break). Figure 8 shows a comparison between active users in ParticipAct during April and August 2014 with the number of active users in MDC Nokia during January and August 2011. It is worth to notice that both of the datasets offer a significant amount of points in crowded and non-crowded time periods. This is particularly evident with

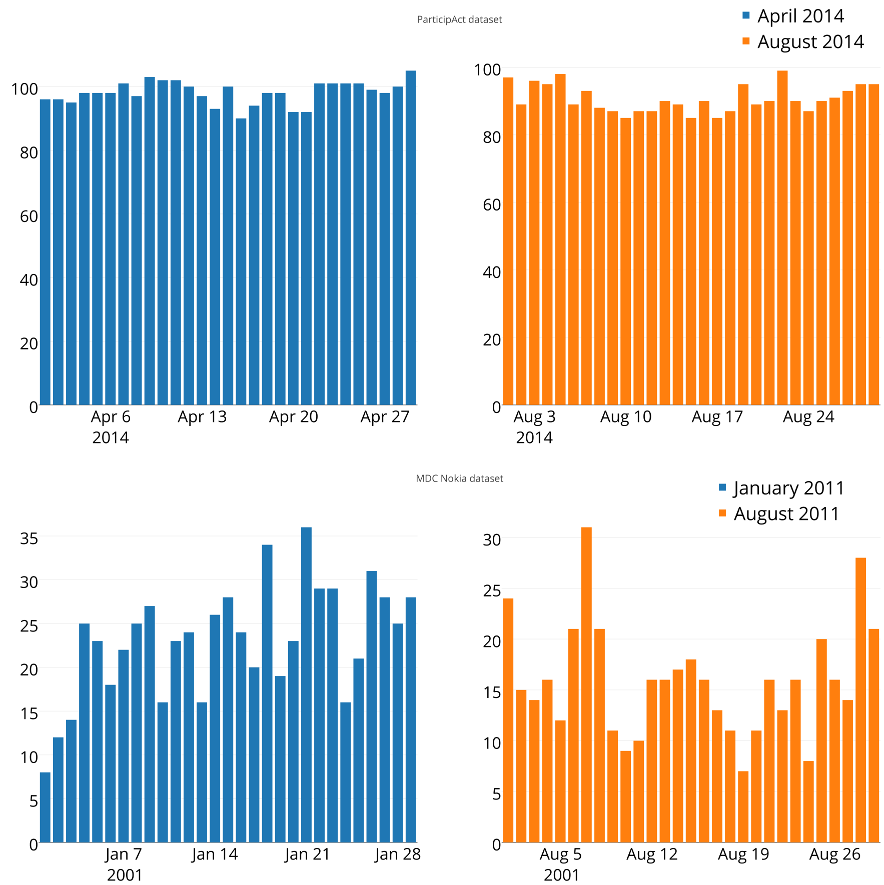


Figure 8: Amount of data points for the ParticipAct and MDC Nokia datasets.

the ParticipAct dataset where the number of available points does not change significantly in different periods. We consider that such rough analysis confirms the capability of the datasets of reproducing a real-world scenario independently from the profiles of users involved in the sensing campaign.

4. MCS Metrics Analysis on the ParticipAct Dataset

In the previous section, we validated the ParticipAct dataset against other similar large datasets in the literature to better define it and to clearly show some of its main characteristics [1, 2, 3, 4, 5]. This section, instead, focuses on the unique characteristics of the ParticipAct dataset and, most importantly, the effectiveness of the management of socio-technical MCS task assignments, currently not supported by other existing living labs available in the literature.

The common goal of all experimental results is to quantify the ParticipAct ability to identify only the users who are more likely to accept and complete a task based on history of their daily movements. This is a core key performance indicator for novel state-of-the-art MCS management platforms because it relevantly impacts on overall MCS campaign efficiently (e.g. completion rate and task execution latency). In fact, ParticipAct task assignment policies focus on geo-executed tasks of special interest for smart city managers to involve in the campaigns the users who are more likely to visit the targeted area and consequently to accept the associated MCS task. In addition, the section provides an in-depth discussion about lessons learned in terms of acceptance, completion rates, and times of assigned tasks (either geo- or non-geo-notified), by considering tasks with different levels of complexity and duration.

4.1. ParticipAct Assignment Policies

ParticipAct enables effective scheduling of geo-executed tasks through four main different policies, specifically designed and optimized for MCS campaigns, namely, random, recency, frequency, and dbscan, as detailed in the following.

The **Random** policy selects a random subset of all available users, regardless of their position history, based on the user ratio parameter, which is defined as the percentage of all available users to be assigned to the task, from 0% up to 100%. It is an uninformed policy and we introduce it as the baseline solution, for the sake of comparison with the other more aware and informed policies.

The **Recency** policy assigns the task to users who have been recently in the geo-execution area. This policy relies on the assumption that those users may return in the same area in their everyday commuting routine. Moreover, the recency policy ranks all potential candidates according to how recently they have been in the geo-execution area, from the most to the least recent. Similarly to the random policy, it may be configured with user ratios from 0% up to 100%, defined as the portion of candidates (starting from higher ranked ones) to select for an active role in the MCS campaign.

The **Frequency** policy assumes that the people who visited more frequently the geo-execution area of the task are the best candidates to select. This policy implicitly assumes that those users usually stay or regularly attend the area. It selects users that have been in the target area in the past and ranks them according to the time that they spent there compared to the time spent in other places. In addition, as for Recency, this policy supports user ratio setting to further limit the number of assigned candidates.

The **Dbscan** policy uses the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to cluster past user location traces [17]. DBSCAN is a density-based clustering algorithm based on the ideas that i)the density of points inside a cluster is much higher than that of the points outside the cluster and ii)the density of points outside a cluster is much lower that the density of any other cluster. DBSCAN has several properties that make it well-suited for MCS task assignment problems: it does not require knowing the number of clusters to be determined a priori, it can detect arbitrarily-shaped clusters, it is robust to noise and outliers, and it is optimized to run on GIS-enabled databases. The dbscan policy runs the DBSCAN algorithm over all past user positions and clusters those users who actually spend a sizeable amount of time in the target geo-execution area as potential candidates. Then, it selects users in a cluster that intersects the geo-execution area. Like all other policies, dbscan allows selecting a proper user ratio setting; however, differently from recency and frequency policies, since dbscan does not provide a ranking, the dbscan policy determines randomly the final set of selected users.

Note that all task assignment policies are sensitive to the size of location history used. As a consequence, in our work we have decided to consider only a limited window of geolocation history of the last days before task start for each user (the default value is two weeks). If we consider all the history of a user, we can select, with higher probability, users who have changed routines and not capable to complete that kind of task. Of course, it is possible to modify that window size or even specify hours of the day and/or days of the week in order to determine, with a finer grained approach, users who can complete tasks in a more specific time window.

Our large-scale ParticipAct experience with in-the-field deployment of MCS campaigns has demonstrated that the choice of the user-to-task assignment policy has a relevant impact on MCS campaign results [3]. Therefore, in the following subsection we discuss some efficiency metrics to assess the performance of assignment policies;these metrics, and the associated assessment results, can be valuable elements in designing and tuning MCS platforms and campaigns.

4.2. Experimental Results

This section first introduces the metrics to compare user assignment policies, then presents a quantitative assessment of our assignment policies for geo-executed tasks, and finally shows a comparison of task acceptance and completion rates/times for different kind of tasks.

4.2.1. MCS Evaluation Metrics

A widespread and internationally recognized consensus on the evaluation metrics for the analysis of MCS campaign performance has still to be reached in the research community, also because of the lack of wide MCS datasets with real-world results about task assignment policies as well as task acceptance and completion. This paper contributes in proposing novel and usable MCS evaluation metrics.

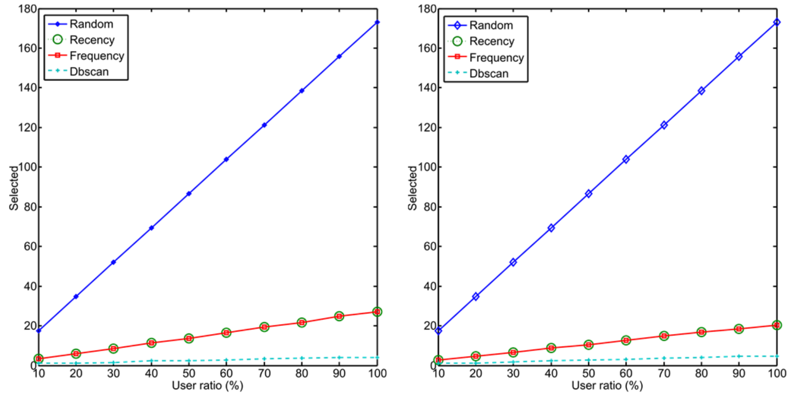
Let us focus first on the MCS metrics to assess the effectiveness of our assignment policies. Assigned users represent the number of candidates to whom different policies assign the task. Precision measures the percentage of success of a given policy, namely, whether selected users actually executed successfully the assigned task. For this evaluation, True Positives (TP) users are the ones who have been selected by a policy and actually carried out the task, while False Positives (FP) are the users selected by a policy and not executing the task. We define precision as the ratio between TP and TP+FP. Moreover, True Negative (TN) users are the ones who have not been selected and did not execute the task, while False Negatives (FN) the users who have not been selected but did execute the task anyway. Accuracy is a percentage and accounts for the proportion of true results (both true positives and true negatives) in the population. More formally, accuracy is the ratio between TP+TN and TP+FP+TN+FN, that quantifies how good is each policy in correctly classifying user behavior and predicting whether they will (TP) or will not (TN) execute a task.

In addition, we introduce some further metrics to evaluate task acceptance and completion. Acceptance rate and completion rate represent the percentages of users (evaluated over all involved people) who, respectively, accepted and completed the task. Similarly, acceptance time and completion time are the times required to, respectively, accept and complete a task.

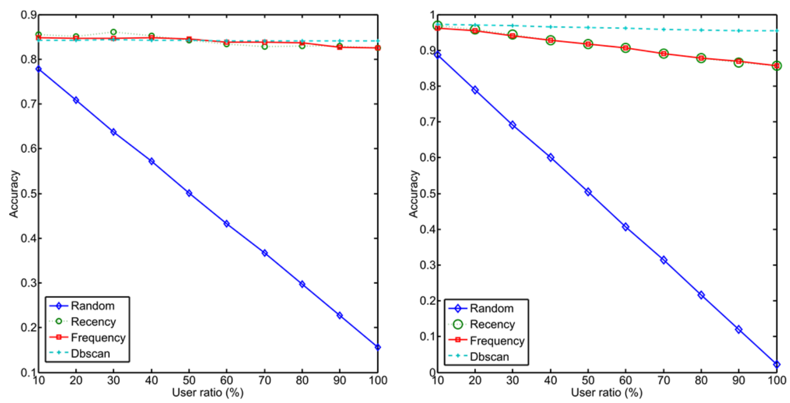
4.2.2. Assignment Policies Evaluation

Figure 9(a) shows the number of candidate users selected by each policy for various geo-executed tasks; we considered 4 geo-notified tasks (graphs on the left) and 4 only geo-executed, but not geo-notified (graphs on the right). For all tasks, we have analyzed the ParticipAct collected data to understand which performance values would have been achieved if users had been selected by different policies and by using different user ratios in the range [10%, 100%]. In the following, all results represent average values over, respectively, the 4 geo-notified and 4 non-geo-notified tasks. Because they were executed by the same population, they have comparable completion/failure rates, and they were associated to urban areas with similar characteristics.

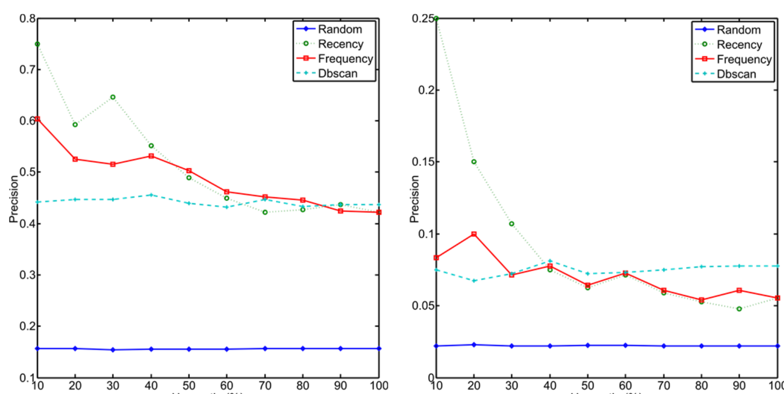
As expected, the random policy shows the worst performance. Recency, frequency, and dbscan policies always select about 20 users or less, with dbscan selecting very few users compared to recency and frequency policies, always below 5 users as average value. These policies (except the random one) significantly limit the number of users assigned to a task but, most relevant, not affecting the final success of the MCS campaign, thus reducing the workload of



(a) Number of candidates selected by each policy (geonotified vs. non-geonotified).



(b) Accuracy on geonotified vs. non-geonotified tasks.



(c) Precision of geonotified vs. non-geonotified tasks.

Figure 9: Policies comparison.

users and the process costs.

Figure 9(b) and Figure 9(c) report the precision and accuracy indicators showing that recency, frequency, and dbscan have strengths and weaknesses of that can indicate their adoption in different application domains and deployment scenarios. Recency and frequency perform similarly and it is important to state that the list of assigned users produced by them contain the same users, but ranked differently: this is why at 100% they reach always the same results. At the same time, let us also rapidly note that some (minor) oscillations in Figure 10 are due to the fact that, notwithstanding the good number of users involved in the ParticipAct MCS campaigns, the cardinality of participants is still sufficiently low to exhibit stochastic fluctuations, especially for recency. Compared to recency and frequency, the dbscan policy has a higher accuracy and shows overall a very stable behavior for all considered metrics. Most important, it obtains those results with a very low number of assigned users, thus confirming dbscan ability to capture and cluster user routinely behaviors.

Finally, a very important lesson learned from our analysis is the geo-notification should be applied whenever possible. Indeed, notifying potential candidates exactly at the time they enter the geo-execution area allows to boost the task completion rate and, consequently, precision. We believe that this beneficial effect is due to the fact that candidates who are notified too in advance tend to forget completing the task as they reach the geo-execution area; for instance, that happens for non-geonotified tasks, with the notification completely detached from the geo-execution, and for geo-notified tasks with a geo-notification area much wider than the geo-execution area and including it.

4.2.3. Task Acceptance and Completion Analysis

Figure 10 reports results on acceptance and completion for different kinds of tasks. Figure 10-a shows the difference in geo-notified (on the left) and non-geo-notified (on the right); considered tasks are geo-executed and similar in terms of geo-execution area dimension and probability of user presence in the considered area. The collected results confirm that geo-notification allows

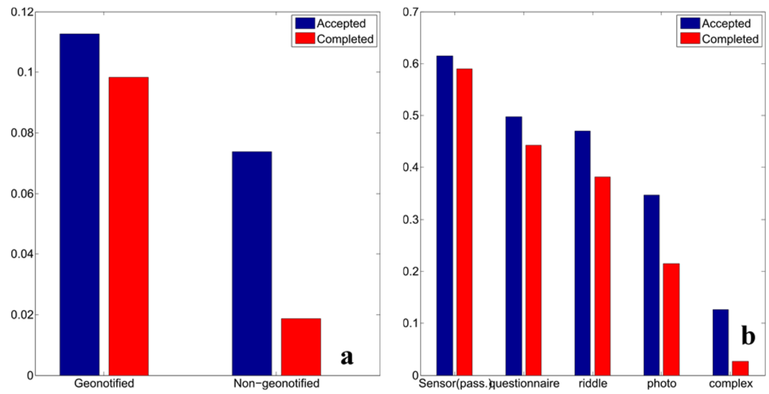


Figure 10: Acceptance and completion rate for different types of task.

improving acceptance and completion rates, respectively, by 52% and 423%.

Figure 10-b, instead, reports acceptance and completion rates for different types of tasks, from passive and easier to complete (short and simple) on the left, to more and more complex ones on the right. Users are more willing to accept and complete passive tasks (without need of explicit user intervention such as GPS monitoring) and simple active tasks (with limited actions). Instead, in case of complex task that aggregate multiple actions, users tend to be less willing to participate and complete them.

We conclude our analysis showing the CCDF for acceptance (Figure 11-a) and completion (Figure 11-b) times depending on geo-notification or not, for the same tasks and users shown in Figure 10. These results are useful to understand and confirm expectations from crowd behaviors. In fact, acceptance time shows to be higher for geo-notified tasks because in this case the user must enter the task notification area to be notified (and then accept the assignment). At the same time, completion time is much less in case of geo-notified tasks because the assigned user is already in the place where to complete the action.

5. Discussion and Lessons Learned

ParticipAct has the main objective of fostering new forms of participation for novel e-citizenship models in the Smart Cities environments and local com-

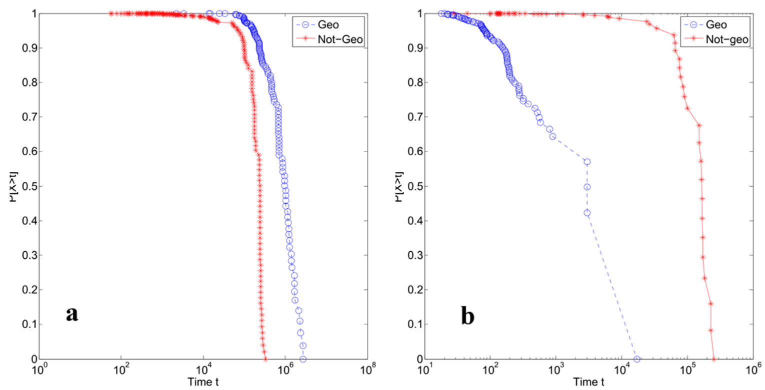


Figure 11: CCDF of acceptance and completion time (s) for geo-notified and non-geo-notified tasks.

munities governance. Our experience with the ParticipAct living lab has given us many insights and allowed us to draw some first conclusions about socio-technical management aspects of MCS, which we believe can be useful to design new campaigns and to refine the whole MCS process.

First of all, let us discuss the characteristics of the ParticipAct dataset compared to the other ones considered in our analysis; Table 1 summarizes the main features of each dataset along four dimensions:

- mobility in terms of frequency of contacts;
- heterogeneity of contacts;
- connectivity in terms of contact duration;
- sensing, which indicates the richness of the sensing data available in the dataset.

As previously discussed, in Cambridge and MIT Reality users tend to meet with many other users for medium/long periods. Instead, in ParticipAct and MDC Nokia (see Table 1) users interact only with a small portion of the whole population. In particular, ParticipAct people meet often but for short periods, and we argue that this dataset well captures the occasional encounters among people in a city. In MDC Nokia people do not meet frequently but the average

Table 1: FEATURES OF EXPERIMENTAL DATASETS

Graphics	***Mobility***	***Heterog.***	***Connectivity***	***Sensing***
Cambridge	high	high	medium	none
MIT	medium	high	high	low
ParticipAct	high	low	low	high
MDC BT	low	low	high	medium
MDC GPS	low+	low+	low+	

contact lasts for longer periods as compared to ParticipAct. Moreover, the exclusive use of GPS coordinates for extracting the co-location traces introduces inaccuracy in the contacts among people, hence also in the final results of the evaluation metrics. Finally, except Cambridge, the other datasets offer some kind of sensing data, but only ParticipAct provides real-world results on user participation to MCS campaigns as thoroughly discussed in Section4.

Second, we analyzed the ParticipAct dataset and the task completion rate during the working (class/studying) hours (i.e., 9-17) and during the rest of the day (i.e., 17-9 off hours) for the different types of tasks. The collected results, depicted in Figure 12, confirm that student participation is well-distributed along the day with an average completion rate of about 50% during working hours and 50% during off hours. Based on these data, we can argue that, after a first period, ParticipAct users tend to consider MCS activities as all their other routinely activities and to schedule them evenly along the day; in its turn, this confirms that the ParticipAct dataset is a realistic living lab able to mimic the use of the smartphone along the whole day for a significant population of users. By focusing on different task types, instead, while sensing activities are typically kept on all day long, more entertaining-related ones (i.e., riddle) are preferably completed during working hours, we believe to take a break, while more complex ones, typically requiring more time and effort, are also preferably completed during the day rather that at late afternoon/night

Third, an important aspect that MCS systems should ascertain and manage

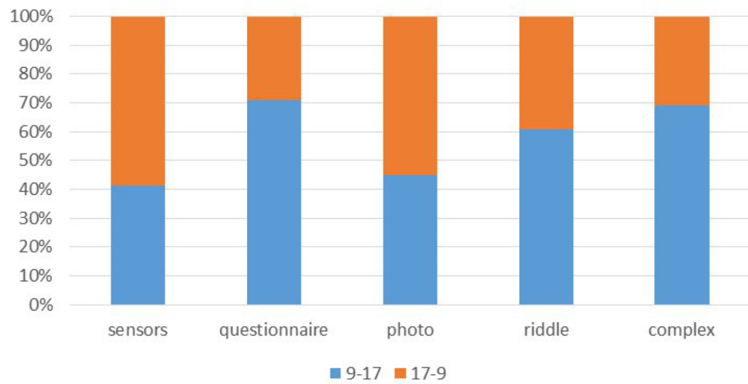


Figure 12: Completion rate of ParticipAct tasks at different time intervals.

is data quality. Crowdsensed data should be refined by keeping into account user trustworthiness (based on her history and reputation) and by enlarging the number of selected users for the same time (to polish data via non-minimal crowdsourcing). For instance, we have observed a minor number of students trying to provide fake data: in most cases, we were able to dynamically detect them for instance when a user completes in a few minutes several tasks that would require taking a photo in places that are several kilometers away; in other situations, only human checking could validate the content, such as when a user, asked to take a picture of a monument, shoots a picture of her monitor that displays the asked monument.

Finally, we found that it is important to make MCS tasks as simple as possible to encourage user participation. In ParticipAct, we have decided to split MCS campaigns into constituent sub-components by assigning independently simple tasks that users can accept more freely. We observed that simple tasks were definitely more easily accepted than complex ones: the distilled guideline is that any MCS system should avoid to ask for big changes in user behavior by soliciting complex tasks that participants are likely to refuse. In particular, we have practically determined as *simple* tasks those ones that are not only easy to understand, but also require little time and minimal physical effort to complete, while not disrupting daily routine. Another important topic is how to minimize user workload and what kind of incentives is more effective in encouraging users to execute even more complex tasks.

6. Conclusive Remarks and Future Work

After some years of intense research in the MCS area, there is still lack of real-world large-scale MCS datasets and Living Labs able to truly verify any step in the whole MCS process, from mobility to task scheduling, from task acceptance to task completion. This paper contributes to fill this gap by presenting a detailed analysis of the ParticipAct MCS dataset. In particular, we originally propose new specific metrics for the analysis of MCS datasets and distill primary lessons learned from the ongoing and large-scale ParticipAct experience, claimed to be useful for researchers and practitioners in the field, to efficiently design new MCS campaigns and to finely tune future crowdsensing processes.

The encouraging results achieved so far are stimulating our further research work, along two primary directions. On the one hand, we are currently extending this research by considering other recent datasets of emerging popularity. In fact, the proliferation of different social networks can be exploited in order to mash-up different sources of information in order to benchmark and integrate data gathered within the ParticipAct platform; we reviewed the Web and the literature of MCS systems and we are currently considering two well-established datasets, namely, Foursquare and TripAdvisor, and the seminal Crowdsignals.io initiative. On the one hand, we are currently extending this research by considering other recent datasets of emerging popularity. In fact, the proliferation of different social networks can be exploited to mash-up different sources of information in order to benchmark and integrate data gathered within the ParticipAct platform. In particular, we are currently considering two well-established datasets, namely, Foursquare and TripAdvisor, and the seminal Crowdsignals.io initiative.

References

[1] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, A. Chaintreau, Crawdad data set cambridge/haggle, (v. 2006-01-31), `http://crawdad.org/cambridge/haggle/` (2006).

[2] N. Eagle, A. Pentland, Reality mining: Sensing complex social systems, Personal and Ubiquitous Computing 10 (4) (2006) 255–268.

[3] G. Cardone, A. Cirri, A. Corradi, L. Foschini, The participact mobile crowd sensing living lab: The testbed for smart cities, IEEE Communications Magazine 52 (10) (2014) 78–85.

[4] N. Kiukkonen, B. J., O. Dousse, D. Gatica-Perez, L. J., Towards rich mobile phone datasets: Lausanne data collection campaign, in: Proc. ACM Int. Conf. on Pervasive Services (ICPS), Berlin, 2010.

[5] J. K. Laurila, D. Gatica-Perez, I. Aad, B. J., O. Bornet, T.-M.-T. Do, O. Dousse, J. Eberle, M. Miettinen, The mobile data challenge: Big data for mobile computing research, in: Pervasive Computing, 2012.

[6] P. Hui, J. Crowcroft, E. Yoneki, Bubble rap: Social-based forwarding in delay-tolerant networks, Mobile Computing, IEEE Transactions on 10 (11) (2011) 1576–1589. `doi:10.1109/TMC.2010.246`.

[7] M. Balazinska, P. Castro, Characterizing mobility and network usage in a corporate wireless local-area network, in: Proceedings of the 1st International Conference on Mobile Systems, Applications and Services, MobiSys '03, ACM, New York, NY, USA, 2003, pp. 303–316. `doi:10.1145/1066116.1066127`.

[8] J. Yeo, D. Kotz, T. Henderson, Crawdad: A community resource for archiving wireless data at dartmouth, SIGCOMM Comput. Commun. Rev. 36 (2) (2006) 21–22. `doi:10.1145/1129582.1129588`.

[9] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, A. T. Campbell, A survey of mobile phone sensing, Comm. Mag. 48 (9) (2010) 140–150. doi:10.1109/MCOM.2010.5560598.

[10] 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.

[11] K. C. . Lin, W. . Lin, C. . Chou, Social-based content diffusion in pocket switched networks, IEEE Transactions on Vehicular Technology 60 (9) (2011) 4539–4548.

[12] A. Passarella, M. Conti, Analysis of individual pair and aggregate inter-contact times in heterogeneous opportunistic networks, IEEE Transactions on Mobile Computing 12 (12) (2013) 2483–2495. doi:10.1109/TMC.2012.213.

[13] A.-D. Nguyen, P. Senac, M. Diaz, Modelling mobile opportunistic networks from mobility to structural and behavioural analysis, Ad Hoc Networks 24, Part B (2015) 161 – 174, modeling and Performance Evaluation of Wireless Ad-Hoc Networks.

[14] T. Karagiannis, J.-Y. Le Boudec, M. Vojnović, Power law and exponential decay of inter contact times between mobile devices, in: Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking, MobiCom '07, ACM, New York, NY, USA, 2007, pp. 183–194. doi:10.1145/1287853.1287875.

[15] A. Passarella, M. Conti, C. Boldrini, R. I. Dunbar, Modelling inter-contact times in social pervasive networks, in: Proceedings of the 14th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, MSWiM '11, ACM, New York, NY, USA, 2011, pp. 333–340. doi:10.1145/2068897.2068955.

[16] F. Ekman, A. Keränen, J. Karvo, J. Ott, Working day movement model, in: Proceedings of the 1st ACM SIGMOBILE Workshop on Mobility Models, MobilityModels '08, ACM, New York, NY, USA, 2008, pp. 33–40. doi:10.1145/1374688.1374695.

[17] M. Ester, H. peter Kriegel, J. S, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: Proc. of 2nd International Conference on Knowledge Discovery and, AAAI Press, 1996, pp. 226–231.

List of Figures