Towards Context of Quality in Mobile Sensing Campaigns

Luis A. Castro
Instituto Tecnológico de Sonora (ITSON)
Ciudad Obregon, Mexico
luis.castro@acm.org

Erica C. Ruiz
Instituto Tecnológico de Sonora (ITSON)
Ciudad Obregon, Mexico
erica.ruiz@itson.edu.mx

Leonardo J. Gutiérrez
Instituto Tecnológico de Sonora (ITSON)
Ciudad Obregon, Mexico
ingleonardogutierrez@hotmail.com

Luis-Felipe Rodríguez
Instituto Tecnológico de Sonora (ITSON)
Ciudad Obregon, Mexico
luis.rodriguez@itson.edu.mx

Abstract
Capturing large amounts of data in mobile sensing campaigns poses serious challenges related to data quality, particularly for advancing research. Data quality processes are usually carried out by researchers after careful scrutiny of data. In this paper, we present a platform for quality assurance in mobile sensing campaigns. The platform aims at working with any mobile sensing platform. It is capable of analyzing and discriminating raw sensor data. In addition, this platform can provide a reputation for the participants based on the quality of the data provided.

Author Keywords
Mobile sensing; Data quality; Mobile sensing campaigns; context of quality

ACM Classification Keywords
H.3.5. Information Storage and Retrieval: Online Information Services

Introduction
Mobile sensing rely on sensor data to capture real world information about users or their surroundings [8]. Mobile sensing campaigns generate a huge quantity of raw sensor data that are mainly complex and heterogeneous [1, 11] due to different vendors, and lack of standards or conventions in data formats. For research, col-

lecting data for mobile sensing can be systematically carried out through research protocols −usually referred to as sensing campaigns− but also in an uncontrolled, realistic fashion through commercial gadgets or with the support of companies (e.g., mobile phone records). While the latter can be used to analyze raw data a posteriori, the former requires careful a priori planning and execution, considering that data are to be used to advance research. In both scenarios, analyzing raw data can be a daunting task due to the process associated to analysis and making sense of the data (e.g., collection, cleaning, labeling, curation, and the like). In this entire process, the quality of the collected data is paramount for obtaining adequate results [1]. Needless to say, for research, the quality of the data is fundamental for obtaining reliable, accurate findings or robust algorithms.

Although important, one of the unattended issues in mobile sensing is that of data quality. Usually, data quality assurance processes are most of the time carried out once the sensing campaigns are complete. Before data analysis, most literature in mobile sensing typically report carrying out a stage called data preprocessing. That is, when all data have been collected, researchers look for noisy, incomplete or atypical data (i.e., outliers) to exclude them from analysis since they tend to bias results or lead to unreliable findings.

In not a few cases, collected data are incomplete, scattered, or inaccurate. From a researcher's perspective, obtaining low-quality data can be frustrating since opportunities for collecting data are at times rare or expensive. Low-quality data can be the result of technologies used, the environments, or user behavior.

Typically, mobile sensing campaigns involve recruiting participants that meet certain eligibility criteria. Participants then follow instructions given by researchers, and often annotate the data that are being collected as per the researcher's request. Sensing campaigns can take from a few days to several months. During this time, it is important to maintain participants well motivated and informed of the many implications of their participations in this process. Since many of these campaigns involve scenarios wherein devices of personal use such as the mobile phone or wearable devices are used, controlling the quality of the data can be difficult due to the changing contexts in which data are generated and captured. As opposed to data generated by static sensors, many participants have a great influence in data quality, due to their changing context, habits, and errors/omissions to the data collection protocol.

When sensing campaigns are for a few weeks long, having the participants behave erratically (e.g., not according to the protocol) can have several implications for data analysis. This is why is important to closely monitor participants' behavior during the sensing campaign. Information about the data contributed by participants would enable the researcher to know in a timely fashion that, for instance, a user usually places her mobile phone on the table rather than carrying with herself, or that the mobile phone has been left with the battery depleted for an entire morning. Therefore, maintain awareness of data quality could enable researchers to discriminate among those participants who may contribute with data that can be ultimately useful.

Context of quality
While most research and development in mobile sensing has focused on developing novel technologies in

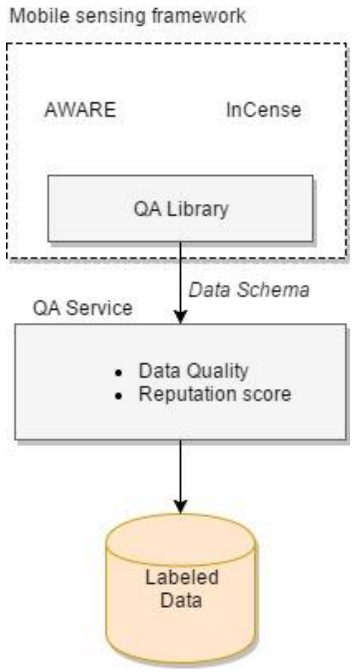

Figure 1:.Schematic of architectural model of the QA service

hardware (e.g., sensor platforms) or software (e.g., mobile sensing frameworks), initiatives for quality assurance (QA) in mobile sensing are exceptional [3, 9]. As mentioned, these initiatives would help researchers to assess the quality of the captured context. Also, researchers could obtain feedback during the data collection process for decision making, reducing costs, and maximizing context quality. Finally, having similar or, better yet, the same metrics for data quality would enable researchers to compare their results with others, or replicate studies within a given community.

Deriving context is per se challenging, but deriving context with low-quality data could yield unreliable context that can ultimately have a negative effect on the efficacy of context-aware applications. Certainly, there are applications that demand deriving context that is complex (e.g., human activity or behavior, internal mental states), while others require much simpler forms of context (e.g., user id). In both cases, as expected, it is fundamental having high-quality data for deriving context of quality.

Our Proposal: A Service for Data Quality

We draw on previous work [2, 4, 6, 10], which mainly involved developing a platform for collecting data from populations of mobile phone users. Based on the literature and our previous work, we account for the need of discriminating data in terms of quality to suggest corrective measures to data or to adjust participants' incentives during the sensing campaign. Given that the data obtained usually pertain to user behavior, in the end, this work aims at having more reliable data obtained from reliable participants.

We are currently working on a platform for data quality, which aims at providing a service for data QA to be used by the research community. The QA service consists mainly of the following modules (see Figure 1):

- Mobile sensing framework (MSF). This can be any sensing framework or platform available in the academic community such as AWARE [5], InCense [2, 4], or Funf .
- QA Service. It is a server-side RESTful service, which will be taking data and the schema of the data as an input, and as an output will be generating a set of vectors describing the data in terms of quality (see [6]) and individuals.
- QA Library. This client-side module works as the communication bridge between the *QA Service* and a *MSF*. Since *QA Service* is RESTful, the implementation of this library can be adapted to the particular language or platform used. It could also implement some on-device analytics for privacy. This module is still work-in-progress.
- Labeled data. Data per se will not be amended, but an alternate vector containing data quality labels will be created describing the analyzed data. Researchers can then use the vectors to discriminate data according to certain quality metrics.

To show the applicability of the proposed platform, we next present an excerpt from a finished sensing campaign, in which the QA platform was used.

Preliminary Evaluation

The Walkability sensing campaign

To show the applicability of the proposed platform, we analyzed data from a completed sensing campaign. We requested access to an anonymized excerpt of a dataset comprising a walkability study, in which they collected data from 65 participants in Mexico, Turkey and Spain regarding their walking habits [7]. The dataset excerpt comprised 354 records pertaining to 29 participants. The purpose of this study was to test the functions used for data analysis. In particular, those related to Data quality and Reputation (see [6] for further details).

Data quality metrics
The platform implements the following metrics: Completeness, timeliness, validity, accuracy, and consistency. However, depending on the purposes of the campaign, and the actual dataset, some of them may not necessarily be used. For this work, Completeness and Validity were used, which are next described:

- Completeness: The proportion of complete stored data. Basically, the way this is measured is checking the absence of blank (null or empty string) values or the presence of non-blank values. This function yields a binary outcome.
- Validity: Data are valid if they conform to the syntax (e.g., format, type, range) of their definition i.e., expected vs. actual value. The measure for this is a comparison between the data and the metadata or documentation for the data item.

For all data metrics, a data schema is needed which basically helps computing data quality metrics according to the expected values. The schema must be provided as an input along with the data.

Instruments
For the purposes of this paper, the experiments were carried out in a computer with the following characteristics: Dell Inspiron 3437, processor Intel Core i5-4200U CPU @ 1.6GHz,, RAM: 4096MB, and 64-bit Windows 7 OS. The DBMS is PostgreSQL (v9.6), and algorithms were implemented in Python (v3.5.0).

Results
In this section, we present the results of using the aforementioned platform in the Walkability dataset. It is important to mention that the results, rather than deriving new knowledge, aim at showing the feasibility of using the platform in collected data.

Following Table 1, the two rightmost columns refer to the Completeness and Validity score per record. A weight *w* can be assigned by the researcher to account for the relevance or importance of a given metric. In this particular case, both were w=0.5. When data are captured by sensors, the completeness value will normally be non-null, but this becomes particularly relevant for users' self-report.

Following Table 2, it can be seen that the reputation score for this particular participant is 84.25 out of 100. Reputation can be updated daily, and can be used to adjust participant incentives. Direct reputation refers to the user's observed behavior i.e., previous data quality records and participants' past performances. On the other hand, indirect reputation refers to the researcher's or the community's trust in the participant. Using reputation, participants can be ranked and rewarded, or the researcher can use it for incentives policies.

ID	Date	Type	Description	Event	Direction	Media	Comment	Completeness (w=.5)	Validity (w=.5)
10	2/20/2016	service	experience	-	res => par	phonecall	we	43.75	43.75
10	2/21/2016	service	technical	-	res => par	email	we	43.75	43.75
10	2/21/2016	question	technical	-	par => res	face2face	-	37.5	37.5
10	2/23/2016	question	technical	-	par => res	face2face	-	37.5	37.5
10	2/24/2016	service	code	2402163K	res => par	email	-	43.75	43.75
							AVG	**41.25**	**41.25**

Table 1: Excerpt of the data corresponding to participants 10; the two rightmost columns were computed from the data using Completeness and Validity metrics.

Reputation	Value
Direct Reputation (w=.9)	74.25
Indirect Reputation (w=.1)	10
TOTAL	**84.25**

Table 2: Reputation score for participant ID=10

Closing remarks

We presented a platform for data quality in mobile sensing campaigns. The results of this work can be used for illustrating how the data quality platform can be used by other researchers interested in having high-quality data in their datasets. Although limited, this work contributes to the mobile sensing community by raising awareness to a pending issue. Certainly, there is room for improvement in terms of performance and scalability that must be addressed before this work can make a real impact.

Although performance is an aspect of interest, and a concern, due to the size of the dataset, it was not possible to carry out a much more comprehensive analysis. Another issue at the moment is that data must be ordered in the same particular order as the schema, which can be solved by the **`key: value`** feature, but this can affect performance since data will have to be looked for. As far as this work is concerned, the biggest performance issue has to do with DB insert, rather than parsing data, but much more comprehensive experiments with actual data are needed. Future work includes investigating factors that deter participants from contributing with high-quality data.

Acknowledgements

This work was partially funded by the National Council for Science and Technology in Mexico (CONACYT) with a scholarship provided to the second author. Also, this work was funded by the Instituto Tecnológico de Sonora (ITSON). We thank Netzahualcóyotl Hernández and Prof. Jesus Favela for providing access to a sample of the Walkabilty sensing campaign.

References

1. Henrik Blunck, Niels Olof Bouvin, Tobias Franke, Kaj Grønbæk, Mikkel B Kjaergaard, Paul Lukowicz and Markus Wüstenberg. 2013. On heterogeneity

in mobile sensing applications aiming at representative data collection. in ACM conference on Pervasive and ubiquitous computing (Ubicomp 2013), ACM, 1087-1098.

2. Luis A. Castro, Jesus Favela, Eduardo Quintana and Moises Perez. 2015. Behavioral data gathering for assessing functional status and health in older adults using mobile phones. Personal and Ubiquitous Computing, 19 (2). 379-391. 10.1007/s00779-014-0825-9
3. Jesus Favela, Luis A. Castro and Layla Michan. 2016. Towards a Federated Repository of Mobile Sensing Datasets for Pervasive Healthcare PervasiveHealth 2016, EAI, Cancun, Mexico.
4. Iván R. Félix, Luis A. Castro, Luis-Felipe Rodríguez and Erica C. Ruíz. 2016. Component-Based Model for On-Device Pre-processing in Mobile Phone Sensing Campaigns. in García, C.R., Caballero-Gil, P., Burmester, M. and Quesada-Arencibia, A. eds. Ubiquitous Computing and Ambient Intelligence: 10th International Conference, UCAmI 2016, Part I, Springer International Publishing, Cham, 201-206.
5. Denzil Ferreira, Vassilis Kostakos and Anind K. Dey. 2015. AWARE: mobile context instrumentation framework. Frontiers in ICT, 2. 6. 10.3389/fict.2015.00006
6. Leonardo J. Gutiérrez, Luis A. Castro, Luis-Felipe Rodríguez and Erica C. Ruiz. 2017. Sensor Data in Internet of Things Applications: A Data Quality Assurance Platform for Research and Development. in 2nd AFI 360° Conference Track on Future Internet and Internet of Things Applications, Toluca, Mexico, EAI.
7. Netzahualcóyotl Hernández, Bert Arnrich, Jesús Favela, Remzi Gökhan, Cem Ersoy, Burcu Demiray and Jesús Fontecha. 2016. mk-sense: An Extensible Platform to Conduct Multi-institutional Mobile Sensing Campaigns. in García, C.R., Caballero-Gil, P., Burmester, M. and Quesada-Arencibia, A. eds. Ubiquitous Computing and Ambient Intelligence: 10th International Conference, UCAmI 2016, San Bartolomé de Tirajana, Gran Canaria, Spain, November 29 – December 2, 2016, Proceedings, Part I, Springer International Publishing, Cham, 207-216.
8. Nicholas D Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury and Andrew T Campbell. 2010. A survey of mobile phone sensing. IEEE Communications magazine, 48 (9).
9. Chris Nugent, Ian Cleland, Anita Santanna, Macarena Espinilla, Jonathan Synnott, Oresti Banos, Jens Lundström, Josef Hallberg and Alberto Calzada. 2016. An initiative for the creation of open datasets within pervasive healthcare. in Proceedings of the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 318-321. 10.4108/eai.16-5-2016.2263830
10. Moises Perez, Luis A. Castro and Jesus Favela. 2011. InCense: A Research Kit to Facilitate Behavioral Data Gathering from Populations of Mobile Phone Users 5th International Symposium of Ubiquitous Computing and Ambient Inteligence (UCAmI 2011), Riviera Maya, Mexico.
11. Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind Dey, Tobias Sonne and Mads Møller Jensen. 2015. Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition. in 13th ACM Conference on Embedded Networked Sensor Systems (SenSys 2015), Seoul, Republic of Korea, ACM, 127-140.