
Towards Altruistic Data Quality Assessment for Mobile Sensing

Eija Ferreira, Denzil Ferreira

Center for Ubiquitous Computing
University of Oulu
Oulu, Finland
eija.ferreira@oulu.fi, denzil.ferreira@oulu.fi

Abstract

High-quality data is a necessity for successful research and development endeavors. In this article, we review relevant literature for data quality (DQ) assessment methods in different domains and discuss the possibilities, challenges and constraints of applying them to mobile sensing. We identify DQ dimensions directly applicable to sensor data: *believability* (comparison with the correct operating bounds), *completeness* (missing values), *free-of-error* (erroneous values), *consistency* (over time), *timeliness* (delay), *accuracy* (deviation from true value) and *precision* (granularity of readings) are core aspects of high-quality sensor data. We also emphasize that sensor data must be *representative* of the originating type of sensor. We propose an altruistic approach to DQ assessment for sensor data that facilitates aggregating and sharing of domain knowledge through a community-contributed library of DQ assessment methods organized by sensor type.

Author Keywords

Mobile instrumentation; Sensors; Data quality; Dimensions; Metrics; Library of domain knowledge.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI); Miscellaneous

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
UbiComp/ISWC'17 Adjunct, September 11–15, 2017, Maui, HI, USA
© 2017 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-5190-4/17/09.
<https://doi.org/10.1145/3123024.3124439>

Current approach:

- Sensor selection
- **Acquiring in-depth knowledge about sensor characteristics**
- **Literature review for DQ assessment methods**
- **Implementation of DQ assessment methods**
- Data collection
- Evaluation of data
- Modeling
- Results

Our approach:

- Sensor selection
- **Library of DQ assessment methods with implementation**
- **Setting minor parameters for the specific sensor model or context of data collection (e.g., sampling rate)**
- Data collection
- Evaluation of data
- Modeling
- Results

Introduction

Smartphones are used to collect massive amounts of data, both through organized data collection efforts for research purposes and by operating systems and mobile applications collecting and storing data for performance and analytics. The range of sensors embedded on the phone (accelerometer, gyroscope, GPS, magnetometer, ambient light, proximity, microphone, *etc.*) and data collected through the operating system APIs (phone and application usage, battery, *etc.*) can be extended by instrumenting the phone with external hardware sensors, asking the user to provide additional contextual information (ESMs [21]) or by retrieving information from online and software sources (*e.g.*, weather, traffic, calendar, email). This data can then be used for various research purposes, such as investigating mobile usage [4], understanding personal [11] and social behavior [2,6], or developing applications for wellness and health promotion [7,14,18] or supporting health care [1].

Irrespective the use of the data, DQ has a major impact on the success of the actions taken based on the data. For research, good DQ is crucial for any data modeling efforts and validity of results. DQ affects the efficiency of interventions and therefore the research's outcome to society. For mobile users, useful, engaging and well-received applications can only exist based on good quality data that truthfully reflects the user's phone usage and behavior (*e.g.*, activities, preferences, routines).

In this article, we overview relevant research and methods for DQ assessment to discuss the possibilities, challenges and constraints of applying such methods to mobile sensing. We start by reviewing previous

definitions of DQ and discussing their suitability for DQ assessment for mobile sensor data. We then propose a modification of an existing definition of DQ which considers the re-occurrence of the same measurement types in datasets collected through smartphones. We also review previously defined DQ dimensions, as well as commonly used metrics for sensor DQ assessment. We discuss their applicability for assessing mobile sensor DQ. Finally, we introduce our approach for DQ assessment in mobile sensing which includes a community-contributed library of DQ assessment methods supporting aggregating and sharing of domain knowledge organized by sensor type.

What is data quality?

DQ has different connotations depending on the field of research or industry. Consequently, there are no standardized cross-application metrics for DQ. In organizational information systems research, DQ refers to how fit the data is for use by data consumers [20]. Sebastian-Coleman [17] defines DQ by two factors: how well does the data meet the expectations of data consumers, and how well does the data represent the objects, events and concepts it is meant to represent. In the field of Internet of Things (IoT), DQ means how suitable the data collected from the smart "things" are for providing ubiquitous services for IoT users [9].

We agree with these general definitions and acknowledge that different aspects of DQ should be considered also in the case of mobile sensing. However, we argue that DQ assessment for mobile data collection could benefit from the fact that it is very often the same type of hardware- (accelerometer, GPS, light, *etc.*), software- (calendar, email, *etc.*) or human-based sensors (user input) [5] that are used when collecting

data through smartphones. Therefore, we define DQ for mobile sensing based on the measurement type (*i.e.*, type of sensor) in question as follows:

How well does the data meet the expectations for the specific type of sensor data?

In other words, we shift the emphasis from an individual user's expectations for the data to expectations of the entire data user community. It is after all the same basic conditions that data originating from a certain measurement type must meet to be suitable for any use. With this definition, we enable sharing and utilization of domain knowledge in assessing DQ for different measurement types.

Data quality assessment

The problem of DQ assessment is commonly addressed through DQ attributes, called *dimensions*. Wang *et al.* [22] categorized the dimensions of high-quality data according to four semantic aspects: *intrinsic*, *accessibility*, *contextual*, and *representational* DQ:

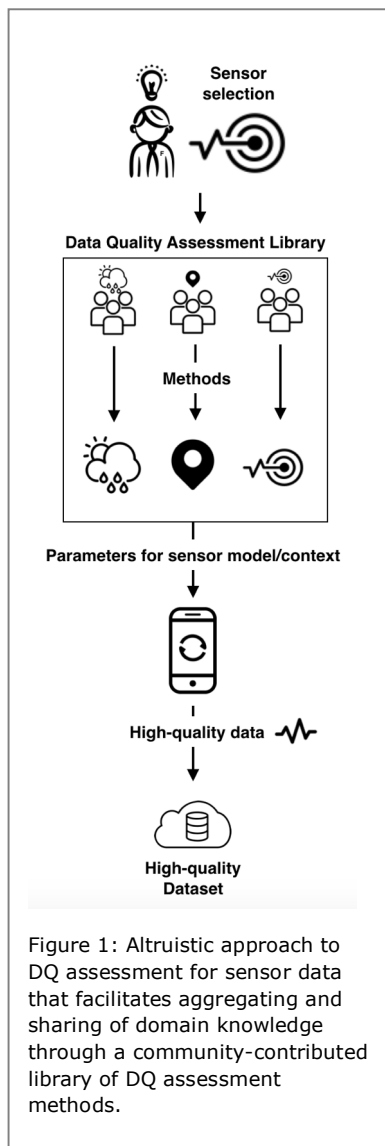
- **Intrinsic** DQ captures the quality that data has on its own (*accuracy, objectivity, believability, reputation*).
- The **accessibility** dimensions of DQ describe how accessible the data is for data consumers (*accessibility, access security*).
- **Contextual** DQ refers to how appropriate the data is for the task at hand (*relevancy, value-added, timeliness, completeness, amount of data*).
- **Representational** DQ describes how understandable and representative the data is (*interpretability, ease of understanding, concise representation, consistent representation*).

In software engineering, the international standard data quality model [8] identifies 15 data quality

characteristics, including *accuracy, completeness, consistency, credibility, currentness, accessibility, compliance, confidentiality, efficiency, precision, traceability, understandability, availability, portability, and recoverability*.

These dimensions have mostly been used in the context of database applications and less work exists to assess quality of real-time data [12]. For sensor data, Klein and Lehner [10] took a step towards automated data quality management by extending the data stream with DQ metadata, using five DQ dimensions: *accuracy, confidence, completeness, data volume* and *timeliness*. In the domain of Wireless Sensor Networks, DQ has been evaluated for example from the perspectives of *timeliness, reliability* and *accuracy* [15]. Karkouch *et al.* [9] survey on techniques of enhancing DQ in IoT highlights the dimensions of *accuracy, confidence, completeness, data volume, timeliness, ease of access, assess security, interpretability, duplicates* and *availability*. In the context of pervasive environments, Li *et al.* [12] use *currency, availability* and *validity*. In the health domain, sensor data is core to medical practice. Sriram *et al.* [19] refer to DQ as the *accuracy, authenticity* and *appropriateness* of a data set for a given purpose.

Commonly, the DQ dimensions are used to derive more specific DQ *metrics* such as *outlier* [9,23] and *anomaly detection* [3]. The problem with the existing metrics is, however, that they are often developed on an *ad-hoc* basis to solve specific problems, and fundamental principles necessary for developing general DQ metrics in practice have not been defined [13].



Various frameworks for DQ assessment have been created based on the DQ dimensions and metrics. Majority of the frameworks are specific to a certain domain such as organizational data management [22], [17] and are intended to be used *post-hoc*, on offline data. Despite the acknowledgement of the importance of DQ assessment in different research areas, a comprehensive approach for assessing DQ for mobile sensing is missing.

Discussion

Data quality for mobile sensing

Assessing DQ for mobile sensing is not trivial, as sensors can be leveraged for many diverse purposes. Sensors' DQ is affected by many factors: limitations of the sensor itself, misplacement, communication errors, power failure, sensor malfunction, human error or intentional misuse. In mobile phone sensing, another challenge for DQ is device heterogeneity and fragmentation (*i.e.*, wide variety of smartphone models, operating systems and software versions available). Unified methods and standards for mobile sensor DQ assessment are needed as measurements from thousands of devices are combined to form larger data sets and applications are expected to function the same independently of the platform in use.

It is also necessary that DQ monitoring for mobile sensing is performed in real-time (*i.e.*, as the data is being collected), to save time and resources in logging, preprocessing, storing and modeling the data that could potentially be flawed. We must perform DQ assessment as an integrated part of the data collection process with unified methods, to enable sharing of data sets and algorithms and to make research results comparable between studies.

Several of the above-mentioned dimensions can be directly applied to DQ assessment for sensor data. Particularly, the dimensions of *believability* (comparison with the correct operating bounds), *completeness* (missing values), *free-of-error* (erroneous values), *consistency* (over time), *timeliness* (delay), *accuracy* (deviation from true value) and *precision* (granularity of readings) are all important aspects of high-quality sensor data. In addition, in line with our definition of DQ for mobile sensing, we emphasize that sensor data needs to be *representative* of the type of sensor it originates from. That is, data collected with multiple sensors of the same type should consistently represent the underlying phenomenon being measured in the same way.

A library of data quality assessment methods

Since it is frequent that the same types of sensor data are collected through smartphones, we can collect DQ assessment methods and algorithms from experts for each type of measurements for sharing with novice users of the sensor. Consequently, we save time from learning in-depth knowledge about each sensor type, as well as searching for and implementing DQ assessment methods for each sensor individually.

To accomplish our vision of aggregated and shareable domain knowledge for sensor DQ assurance, we propose to create a community-contributed library of DQ assessment methods organized by sensor type (Figure 1). The library must be open-sourced (*e.g.*, Apache 2.0) and invite contributions from anyone working with sensor data. To facilitate sharing and accessibility, the DQ assessment methods can be implemented with free software such as R [16], which offers a wide variety of packages with readily available

algorithms and functions and integrates easily with any data collection platform (e.g., AWARE [5]). Over time, such library would include a wide range of DQ assessment methods for different types of sensors used in mobile sensing accessible to everyone.

Conclusion

In this article, we have identified an opportunity to benefit from sharing of domain knowledge and methods for DQ assessment in mobile sensing. For the development of future mobile ubiquitous systems and IoT applications, we argue that we should not focus on mere quantity of sources of data, but instead on the quality of the sensors and the data they produce.

We believe that a unified DQ assessment library will be useful, not just for smartphones' sensor data campaigns, but for any sensor data collection effort. The DQ dimensions applicable to sensor data identified in this article can be extended in future work. However, our article's main contribution is a call for open discussion on what strategies fellow scientists employ today to minimize data loss, and supervise data quality. Such discussion across disciplines allows us to better understand the users' expectations for the data and whether it truly represents what we wish to measure with the sensor.

Acknowledgements

This work is partially funded by the Academy of Finland (Grants 276786-AWARE, 286386-CPDSS, 285459-iSCIENCE, 304925-CARE), the European Commission (Grant 6AIKA-A71143-AKAI), and Marie Skłodowska-Curie Actions (645706-GRAGE).

References

1. Sangwon Bae, Denzil Ferreira, Brian Suffoletto, Juan C. Puyana, Ryan Kurtz, Tammy Chung, and Anind K. Dey. 2017. Detecting Drinking Episodes in Young Adults Using Smartphone-based Sensors. *IMWUT*. 1, 2, Article 5, DOI: <https://doi.org/10.1145/3090051>
2. Anabela Berenguer, Jorge Goncalves, Simo Hosio, Denzil Ferreira, Theodoros Anagnostopoulos, Vassilis Kostakos. 2017. Are Smartphones Ubiquitous?: An in-depth survey of smartphone adoption by seniors. *IEEE Consum. Electron. Mag.* 6, 1: 104-110. DOI: <http://dx.doi.org/10.1109/MCE.2016.2614524>
3. Michael W. Bigrigg, H. Scott Matthews, and James H. Garrett, Jr. 2010. Fault perturbations in building sensor network data streams. *Int. J. Sen. Netw.* 7, 3: 152-161. DOI: <http://dx.doi.org/10.1504/IJSNET.2010.033117>
4. Karen Church, Denzil Ferreira, Nikola Banovic, and Kent Lyons. 2015. Understanding the Challenges of Mobile Phone Usage Data. In *Proc MobileHCI '15*. 504-514. DOI: <http://dx.doi.org/10.1145/2785830.2785891>
5. Denzil Ferreira, Vassilis Kostakos and Anind K. Dey 2015. AWARE: Mobile Context Instrumentation Framework. *Frontiers in ICT* 2, 6: 6. DOI: <http://dx.doi.org/10.3389/fict.2015.00006>
6. Denzil Ferreira, Vassilis Kostakos and Immanuel Schweizer. 2017. *Participatory Sensing, Opinions and Collective Awareness*. Springer.
7. Tian Hao, Guoliang Xing, and Gang Zhou. 2013. iSleep: Unobtrusive Sleep Quality Monitoring Using Smartphones. In *Proc ACM Sensys'13*, ACM, 4:1-4:14. DOI: <http://dx.doi.org/10.1145/2517351.2517359>
8. International Organization for Standardization. 2008. *Software engineering - Software product Quality Requirements and Evaluation (SQuARE) - Data quality model*.
9. Aimad Karkouch, Hajar Mousannif, Hassan Al Moatassime, and Thomas Noel. 2016. Data quality

- in internet of things. *J. Netw. Comput. Appl.* 73, C, 57-81. DOI:
<https://doi.org/10.1016/j.jnca.2016.08.002>
10. A. Klein and W. Lehner. 2009. Representing Data Quality in Sensor Data Streaming Environments. *J. Data and Information Quality* 1, 2, Article 10. DOI:
<https://doi.org/10.1145/1577840.1577845>
11. Ian Li, Anind K. Dey, and Jodi Forlizzi. 2011. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proc UBICOMP'11*, 405-414.
<http://dx.doi.org/10.1145/2030112.2030166>
12. Fei Li, Stefan Nastic, and Schahram Dustdar. 2012. Data quality observation in pervasive environments. In *Proc CSE'12*. 602-609.
13. Leo L. Pipino, Yang W. Lee, and Richard Y. Wang. 2002. Data quality assessment. *Commun. ACM* 45, 4: 211-218. DOI:
<http://dx.doi.org/10.1145/505248.506010>
14. Riitta Pyky, Heli Koivumaa-Honkanen, Anna-Maiju Leinonen, Riikka Ahola, Noora Hirvonen, Heidi Enwald, Tim Luoto, Eija Ferreira, Tiina M. Ikäheimo, Sirkka Keinänen-Kiukaanniemi, Matti Mäntysaari, Timo Jämsä, and Raija Korpelainen. 2017. Effect of tailored, gamified, mobile physical activity intervention on life satisfaction and self-rated health in young adolescent men: A population-based, randomized controlled trial (MOPO study). *Comput. Hum. Behav.* 72, C, 13-22. DOI:
<https://doi.org/10.1016/j.chb.2017.02.032>
15. Zhijing Qin, Qi Han, Sharad Mehrotra, and Nalini Venkatasubramanian. 2014. Quality-Aware Sensor Data Management. In: Ammari H. (eds) *The Art of Wireless Sensor Networks. Signals and Communication Technology*. Springer, Berlin, Heidelberg
16. R: The R Project for Statistical Computing. *R: The R Project for Statistical Computing*. <http://www.r-project.org>.
17. Laura Sebastian-Coleman. 2012. *Measuring Data Quality for Ongoing Improvement: A Data Quality Assessment Framework*. Newnes.
18. Pekka Siirtola, Heli Koskimäki, and Juha Röning. 2016. From user-independent to personal human activity recognition models using smartphone sensors. In *Proc ESANN'16*, 471-476.
19. Janani Sriram, Minh Shin, David Kotz, Anand Rajan, Manoj Sastry, and Mark Yarvis. 2009. Challenges in Data Quality Assurance in Pervasive Health Monitoring Systems. In *Future of Trust in Computing* (eds.). Vieweg+Teubner, 129-142.
20. Diane M. Strong, Yang W. Lee, and Richard Y. Wang. 1997. Data quality in context. *Commun. ACM* 40, 5: 103-110. DOI:
<http://dx.doi.org/10.1145/253769.253804>
21. Karel Vandenbroucke, Denzil Ferreira, Jorge Goncalves, Vassilis Kostakos, and Katrien De Moor. 2014. Mobile cloud storage: a contextual experience. In *Proc MobileHCI '14*. 101-110. DOI:
<http://dx.doi.org/10.1145/2628363.2628386>
22. Richard Y. Wang and Diane M. Strong. 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. *J. Manage. Inf. Syst.* 12, 4: 5-33. DOI:
<http://dx.doi.org/10.1080/07421222.1996.11518099>
23. Yang Zhang, Nirvana Meratnia, and Paul J. M. Havinga. 2010. Ensuring High Sensor Data Quality Through Use of Online Outlier Detection Techniques. *Int. J. Sen. Netw.* 7, 3: 141-151. DOI:
<http://dx.doi.org/10.1504/IJSNET.2010.033116>