Activity Self-Tracking with Smart Phones: How to Approach Odd Measurements?

Gabriela Villalobos-Zúñiga
University of Lausanne
gabriela.villalobos@unil.ch

Mauro Cherubini
University of Lausanne
mauro.cherubini@unil.ch

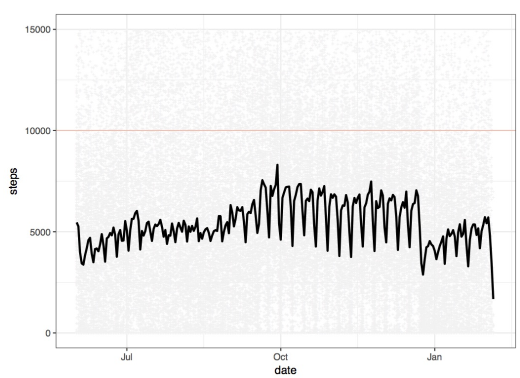

Figure 1: Seasonality Effect - General Curve of Means. The "pulse" regularity is an indication of seasonality effects.

Abstract

Tracking physical activity reliably is becoming central to many research efforts. In the last years specialized hardware has been proposed to measure movement. However, asking study participants to carry additional devices has drawbacks. We focus on using mobile devices as motion sensors. In the paper we detail several issues that we found while using this technique in a longitudinal study involving hundreds of participants for several months. We hope to sparkle a lively discussion at the workshop and attract interest in this method from other researchers.

Author Keywords

Long-Term Self-Tracking; Health; Physical Activity; Longitudinal study; Apple HealthKit

ACM Classification Keywords

H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction

Physical inactivity is the fourth leading risk factor for global mortality [1]. Diseases such as coronary heart disease, type 2 diabetes, breast and colon cancers are caused by insufficient physical activity. Furthermore, the lack of moderate

[1] http://www.who.int/topics/physical_activity/en/,last retrieve February 2018.

Statistic	Value
Min.	0
1st Qu.	2765
Median	5584
Mean	6490
3rd Qu.	8916
Max.	98065
NA's	1760

Table 1: Descriptive statistics of the step count distribution.

intensity physical activity, is responsible for 9% of premature mortality and 3.2 million deaths globally [11]. Due to the overwhelming scientific evidence on the benefits of physical activity [17], it becomes clear the necessity to implement mechanisms to increase the physical activity levels worldwide. Guidelines recommend adults to practice a moderate-intensity physical activity for at least 30 minutes a day during 5 days a week [5].

The points raised before justify the growing body of research aiming to analyze, support and/or enhance human activities through the means of technology [4],[2],[9],[13]. These research efforts focus on diverse populations: diabetics [1], cancer survivors [16],[7], children [15], etc. being their common goal to attain the adoption and/or maintenance of physical activity. Furthermore, industry is as well engaged in the implementation of technology supported means to stimulate physical activity. Examples of these include, fitness coaching apps[2], and run tracking applications[3].

From a hardware standpoint, there has been a growing number of specialized devices that have been developed to sense human activity. These include bracelets, clips[4] or wearable sensors[5]. While as consumer products these devices are great, as research devices they suffer serious limitations. For instance, users might forget to wear them consistently, or their battery could run out of power if not systematically recharged. These points –among others– persuaded us to reconsider employing just smart phones without any additional hardware as a tool for tracking steps.

[2]E.g., https://www.skimble.com/, last retrieved February 2018.

[3]Eg., https://www.runtastic.com/, last retrieved February 2018.

[4]Eg., https://www.fitbit.com/zip, last retrieved February 2018

[5]Eg., https://www.dexcom.com/continuous-glucose-monitoring, last retrieved February 2018.

Using Smart Phones To Track Steps

An elegant solution to these limitations is that of using smart phones as sensors of physical activity. These can accurately predict walking activity [14] and have been previously used to measure physical exercise [12],[8]. Moreover, they have the following advantages: they act as *silent observers*, letting participants carry on with their tasks without an explicit reminder of being tracked, thus making the data capturing less intrusive. Furthermore, people are willing to carry their phones with them during the majority of their daily activities and are also aware of the level of battery these have because they want to remain active in their social networks (e.g., Facebook, Snapchat, Twitter, Instagram) and reachable to instant messaging and calls (e.g., Whatsapp, Skype, SMS). The fact that most of day-time period users carry their phones, diminishes the negative factors that come naturally with the usage of trackers where users forget to charge it or wear it. Finally, another methodological advantage of using the smart phone as a sensor is that it does not require participants to use additional hardware that they usually would not, thus increasing the ecological validity of the study.

Of the different hardware solutions, we focus on the iPhone because Apple has standardized both the hardware and the API with which we can collect activity data, furthermore its use diminishes the development costs required to build the experimental infrastructure for a study.

Apple's HealthKit™ is a platform with a repository of physical activity data collected from the iPhone's accelerometer and other health data obtained from various sensors such as scales and blood testing devices. Data collected is stored in an encrypted database called *Healthkit Store*

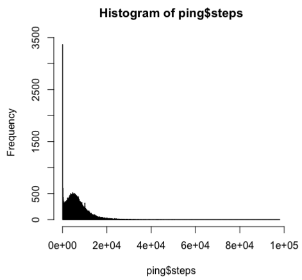

Figure 2: Step Count Distribution. The distribution is left skewed due to the presence of small steps measurements returned by the smart phone.

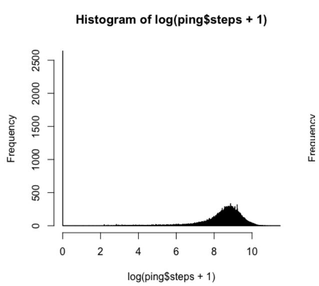

Figure 3: Log-transform of the step count distribution. While the small values are still visible, the distribution assumes a more symmetrical shape.

from which steps are retrieved[6]. Researchers have made use of this platform in the past to implement steps tracking applications [3],[10],[18].

As part of ongoing research we used Apple's HealthKit™ platform as a sensor to collect physical activity data. The goal of this paper is to discuss advantages and disadvantages of this method, and to report specific issues we encountered and possible solutions. We plan to contribute to the workshop by sparkling interest in this method and discussion on how to address some of the issues that we encountered.

Adversities of Tracking Steps With a Smart Phone

We are currently using Apple's Health Kit™ to record the physical activity of users as part of a research project. The study involves an longitudinal observation period of 6 months. 230 students are currently participating in this research(62% female and 38% male), average age of 21(SD = 2.3). Data capture started in June 2017 and it is still ongoing. While analyzing the data, we observed a number of issues that we would like to discuss here.

Unreliability of the sensor to very low physical activity: We observed a conspicuous number of days for which the activity sensors provided unreliable data. These are days which correspond to extremely low physical activity of the participants. During those days, the sensor was unable to provide an accurate measure of the activity of the participants providing counts equal to zero or little steps or returning a NA (i.e., not available) altogether. Obviously, this does not correspond to reality as we interviewed participants and indeed they moved during these days. See Table 1 for descriptive statistics of the steps distribution. We are still unsure of why this happens. This could be due to the inability of the internal accelerometer and its corresponding signal processing algorithm to distinguish human steps when the pace is shorter or slower than usual (e.g., moving indoor vs. outdoor). **Solution**: Our short-term solution to this problem is that of discarding low activity values. To do that we plotted a histogram of the collected steps and noticed skewness towards zero of the distribution curve. We took the logarithmic-transform of the curve and noticed that to make it symmetrical we should remove measurements below 400 steps. See Figures 2 and 3 for context.

Health App configuration changes: In order for Health Kit™ to record steps, the user needs to grant permission to access sensor data. This is is verified by turning on the *Fitness Tracking* switch on the iPhone menu *Settings->Privacy->Motion & Fitness*. In our study we noticed periods in which despite of receiving normal data since the beginning of the study we suddenly started obtaining zeros, see Figure 4. When contacting participants and verifying their iPhone configuration we noticed that the Fitness Tracking switch was turned off, resulting in the inability of the sensor to register steps. Another similar issue occurred when the participant did not grant our research application the rights to access Health data, from which we retrieve the number of steps. **Solution**: Our practical solution was to contact directly these participants and follow a step by step procedure to verify that all the configuration parameters were set up accordingly. We also implemented a feature in the research application that checked the settings whenever the application was opened. If these were turned off, the user was redirected to the *Motion & Fitness* settings to turn the switch back on.

Micro-holes in the dataset: We observed an evident number of days in which the count of steps was reported zero.

[6] HealthKit also collects data from a variety of other sensors. However, in the context of this paper we focus only on steps measurements.

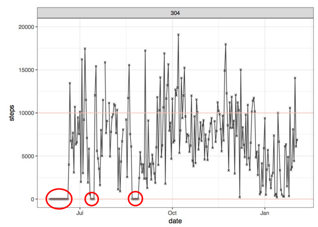

Figure 4: Example of micro-holes in the longitudinal measurements (circled in red). The number on the top of the graph represent the participant id.

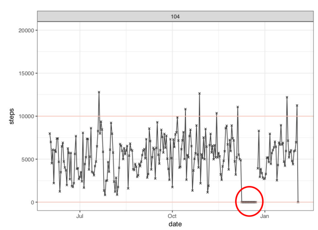

Figure 5: Example of macro-holes.

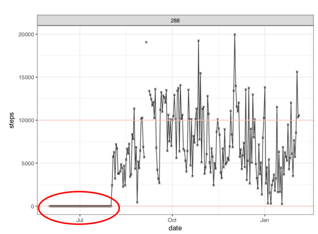

Figure 6: Additional example of macro-holes.

This occurs randomly across the dataset. See Figure 4. We believe this appears due to a technical error in the sensor that prevents it from recording steps. **Solution**: Our solution approach is the same as the one employed for unreliability of the sensor to very low physical activity.

Macro-holes in the dataset: There is no activity recorded for several consecutive days. We believe that these 'macro-holes' are generated in two distinct situations: in some cases, we observe continuous zeros for some participants at the onset of the study. The justification for this is that the research application was not yet installed into the participant's device, therefore no permission to access sensor data was granted. In some other cases, if the macro-hole is found in the middle of the study, this might be caused by the participant who turned off the permission for our app to access the sensor data. See Figures 5 and 6. **Solution**: The approach here is to use a statistical analysis method that is able to deal with missing data. We chose to use Linear Mixed-Effects Regression (or LMER). Each subject in these LMER models may vary in terms of the number of measurement occasions. Subjects who are missing data at a given time point are not excluded from the analysis. In considering missing data and whether they are ignorable or not, a related issue is the distinction between *attrition* (i.e., subjects dropping out of the study and not returning) and *sporadic* or *intermittent data* (i.e., subjects with missing data between observed time-points). Attrition should be consider carefully and might not be ignored. Participants where consistent attrition is observed should be removed from the analysis.

Plateau in the dataset: We noticed data for which in a particular moment in time augmented until reaching above 10'000 steps a day and remained constant for some months. We believe participants in this situation were trying to reach the daily step goal we set to 10'000 steps. See Figure 7. **Solution**: It is still not clear in our analysis whether these participants cheated by tampering the sensor or whether they actually had physical activity consistently above the set threshold of 10K steps for several weeks in a row. To counter people who wanted to cheat in the experiment, we explicitly excluded manually entered steps from the step count. However, participants might have found new creative ways to cheat that we have not identified yet.

Seasonality effects: Physical activity of the participants was influenced by the summer holiday season (July and August). Period in which participants tend not to carry their phone or they perform more or less exercise than their normal average. Similarly, occurs during weekdays and weekends in which people tend to be less active in the later one. These effect creates uncertainty in the effects of the treatment of the study by not allowing to clearly determine what increased or decreased the physical activity. See Figure 1. **Solution**: Our solution for this case consists in removing the seasonality effects by putting in practice methods borrowed from time-series analysis such as effect decomposition.

Discussion

Determining the correct type of devices to track data for research studies is critical for their validity. This decision might depend on the design of the experiment and the specific research questions the researchers are tackling. In some cases it might be adequate to ask participants to carry additional hardware, in others it might not. We argue that even with the limitations presented in this position paper, using smart phones as sensor of physical activity is a powerful mechanism to study human activity and a valid research method. There are a number of open questions that we would like to ask the audience and get feedback on during the workshop.

Which method should we use to process micro-holes in the dataset? There are two possible solutions to this: either we use the last-value carried forward (or LVCF) or interpolation between two known measurements. While the former is more respectful of the auto-regressive nature of longitudinal data [6], the latter is not. We would like to hear expert opinions on this issue.

What are the possible explanations for NAs being returned from Apple HealthKit™? We know little of the inner working of the sensor and the built-in algorithm that process the accelerometer signals in order to understand the fringe cases that we observed. We would like to ask the audience whether any technical documentation is available on this specific hardware or any engineer who might have reverse-engineered the inner working of this black box.

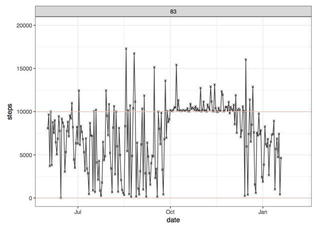

Figure 7: Plateau effect, as observed in one of the participants.

Removing seasonality effects or keeping them in the data? We are still unsure whether the best method to analyze longitudinal data is that of removing the seasonality effects. Contrary to time series, where measurements are said to be independent, data coming from longitudinal studies have an auto-regressive nature that we should not ignore. Seasonality is certainly part of how people behave: activity on week days might be consistently lower for many users. Should we remove this effects?

How to best exploit iPhone background execution mode for experimental purposes? Apple allows background mode for specific purposes such as: playing audio, receiving location updates, performing finite-length task, and background fetch (e.g. news retrieval). Unfortunately none of those fit out need, and this forced us to ask our participants to open the application periodically (bring it to foreground mode) so we could retrieve the steps from HealthKit™ . Avoiding intervening so directly into the experiment, would increase the ecological validity of the study. Recently, we found a mechanism in which by using the option : *UIApplicationBackgroundFetchIntervalMinimum* we are able to overcome this issue. Nonetheless, we would love to discuss whether other solutions to this problematic have been exploited successfully.

Conclusion and Next Steps

Having clear/reliable data is central to long-term self tracking, hence the importance of working with precise and solid measure instruments. Researchers can make use of the variety of trackers that are offered by the market, however all existing solutions have drawbacks. With the purpose of obtaining valid results, it is imperative that researchers find workarounds to standing issues, making the discussion of these problems highly relevant for the research community.

In a follow-up study we are planning to improve the quality of the data we capture working on the different aspects discussed in this paper.

REFERENCES

1. Stephanie E Bonn, Christina Alexandrou, Kristin Hjörleifsdottir Steiner, Klara Wiklander, Claes-Göran Östenson, Marie Löf, and Ylva Trolle Lagerros. 2018. App-technology to increase physical activity among patients with diabetes type 2-the DiaCert-study, a randomized controlled trial. *BMC public health* 18, 1 (2018), 119.

2. François Boudreau, Michel Moreau, and José Côté. 2016. Effectiveness of computer tailoring versus peer support web-based interventions in promoting physical activity among insufficiently active Canadian adults with type 2 diabetes: protocol for a randomized controlled trial. *JMIR research protocols* 5, 1 (2016).

3. Julian Frank. 2017. *A personalized support tool for the training of mindful walking: The mobile 'MindfulWalk' application*. Ph.D. Dissertation. Ulm University.

4. Stijn AH Friederichs, Anke Oenema, Catherine Bolman, Janneke Guyaux, Hilde M van Keulen, and Lilian Lechner. 2014. I Move: systematic development of a web-based computer tailored physical activity intervention, based on motivational interviewing and self-determination theory. *BMC public health* 14, 1 (2014), 212.

5. Carol Ewing Garber, Bryan Blissmer, Michael R Deschenes, Barry A Franklin, Michael J Lamonte, I-Min Lee, David C Nieman, and David P Swain. 2011. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Medicine and science in sports and exercise* 43, 7 (2011), 1334–1359.

6. Ramazan Gençay, Michel Dacorogna, Ulrich A Muller, Olivier Pictet, and Richard Olsen. 2001. *An introduction to high-frequency finance*. Elsevier.

7. Ashley F Haggerty, Andrea Hagemann, Matthew Barnett, Mark Thornquist, Marian L Neuhouser, Neil Horowitz, Graham A Colditz, David B Sarwer, Emily M Ko, and Kelly C Allison. 2017. A Randomized, Controlled, Multicenter Study of Technology-Based Weight Loss Interventions among Endometrial Cancer Survivors. *Obesity* 25, S2 (2017).

8. Tim Harries, Parisa Eslambolchilar, Ruth Rettie, Chris Stride, Simon Walton, and Hugo C van Woerden. 2016. Effectiveness of a smartphone app in increasing physical activity amongst male adults: a randomised controlled trial. *BMC public health* 16, 1 (2016), 925.

9. Robert Hurling, Michael Catt, Marco De Boni, Bruce William Fairley, Tina Hurst, Peter Murray, Alannah Richardson, and Jaspreet Singh Sodhi. 2007. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. *Journal of medical Internet research* 9, 2 (2007).

10. Hayeon Jeong, Heepyung Kim, Rihun Kim, Uichin Lee, and Yong Jeong. 2017. Smartwatch Wearing Behavior Analysis: A Longitudinal Study. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 60.

11. I-Min Lee, Eric J Shiroma, Felipe Lobelo, Pekka Puska, Steven N Blair, Peter T Katzmarzyk, Lancet Physical Activity Series Working Group, and others. 2012. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The lancet* 380, 9838 (2012), 219–229.

12. Chung-Tse Liu and Chia-Tai Chan. 2016. Exercise Performance Measurement with Smartphone Embedded Sensor for Well-Being Management. *International journal of environmental research and public health* 13, 10 (2016), 1001.

13. Bjarne M Nes, Christian R Gutvik, Carl J Lavie, Javaid Nauman, and Ulrik Wisløff. 2017. Personalized activity intelligence (PAI) for prevention of cardiovascular disease and promotion of physical activity. *The American journal of medicine* 130, 3 (2017), 328–336.

14. Meaghan Nolan, J Ross Mitchell, and Patricia K Doyle-Baker. 2014. Validity of the Apple iPhone®/iPod Touch® as an accelerometer-based physical activity monitor: a proof-of-concept study. *Journal of Physical Activity and Health* 11, 4 (2014), 759–769.

15. Anni Pakarinen, Heidi Parisod, Jouni Smed, and Sanna Salanterä. 2017. Health game interventions to enhance physical activity self-efficacy of children: a quantitative systematic review. *Journal of advanced nursing* 73, 4 (2017), 794–811.

16. Michael C Robertson, Edward Tsai, Elizabeth J Lyons, Sanjana Srinivasan, Maria C Swartz, Miranda L Baum, and Karen M Basen-Engquist. 2017. Mobile health physical activity intervention preferences in cancer survivors: a qualitative study. *JMIR mHealth and uHealth* 5, 1 (2017).

17. Darren ER Warburton, Crystal Whitney Nicol, and Shannon SD Bredin. 2006. Health benefits of physical activity: the evidence. *Canadian medical association journal* 174, 6 (2006), 801–809.

18. Tongtong Xu, Ao Guo, and Jianhua Ma. 2017. Analysis of Temporal Features in Data Streams from Multiple Wearable Devices. In *Cybernetics (CYBCONF), 2017 3rd IEEE International Conference on*. IEEE, 1–6.