Using Smartphones for Research Outside Clinical Settings: How Operating Systems, App Developers, and Users Determine Geolocation Data Quality in mHealth Studies

AL Beukenhorst, DM Schultz, J McBeth, R Lakshminarayana, JC Sergeant, and WG Dixon

Abstract

Smartphones that collect user geolocation provide opportunities for mobile Health (mHealth). Although granularity of geolocation data may be high, data completeness depends on the device’s operating system, application developer decisions, and user actions. We investigate completeness of geolocation data collected via smartphones of 5601 people that self-reported daily chronic pain symptoms on 349,293 days. On 17% of these days, hourly geolocation data is reported, but days with 0 (16%), 1 (14%), and 2 (13%) geolocations are common. Android phones collect geolocation more often than iPhones (median 17 versus 2 times a day). Factors on operating system level and individual user level influence completeness of geolocation data collected with smartphones. mHealth researchers should be aware of these factors when designing their studies. The mHealth research community should devise standards for reporting geolocation data quality, analysing systematic differences in data quality between participant groups, and methods for data imputation.

Keywords:
Mobile Health (mHealth); Smartphone; Longitudinal Studies

Introduction

The high penetration of smartphones in our society enables large-scale collection of self-reported data in real-time for research outside clinical settings. Moreover, these temporally-rich data sets can be complemented with passively collected data by smartphone sensors, such as physical activity, sleeping patterns or geolocation data [1]. Geolocation data is especially useful as it can be linked to openly available data sets with high spatial and temporal resolution (e.g. climate, weather, air pollution).

Usage of smartphones to collect geolocation data passively has many advantages. It reduces the burden of participation as participants can use their own device and do not need to manually enter location data. It is less prone to recall bias compared to keeping location diaries, increasing the internal validity of studies [2]. Through the Global Positioning System (GPS), location can be sampled both frequently and accurately, rendering assumptions that participants always stay at their home or work postcode unnecessary.

Although granularity of geolocation data may be higher using smartphones, completeness and accuracy of passively collected sensor data may be less when using smartphones rather than clinically-approved devices and more variable between participants who use different smartphones. The accuracy and completeness of geolocation data are affected by a variety of factors relating to operating system [3], application (app) developer decisions [4; 5] and smartphone settings defined by the user [4; 5], which are partially outside researchers’ control. Thus, if geolocation completeness patterns systematically differ between participant groups, the validity of study results may be compromised.

The purpose of this study is to assess completeness of geolocation data collected during the study Cloudy with a Chance of Pain. This is a national UK smartphone study investigating the association between the weather and chronic pain. Study participants rate ten aspects of their daily symptoms in the study’s smartphone app, which collects geolocation data to retrieve hourly weather data from the nearest weather station via the Met Office DataPoint service.

We investigated data completeness for the study population and compared the number of geolocation observations between Android and iPhone users, as well as their demographics.

Sampling geolocation data: location services on Android and iOS smartphones

The availability of geolocation data depends on the smartphone’s operating system, various settings defined by the app developer, and phone settings chosen by the user [4; 5]. The “Cloudy with a Chance of Pain” app was developed for the two operating systems with largest market share in the UK: Android (53% market share), developed by Google and used on many different brands of phones, and iOS (44% market share), developed by Apple for iPhones only. We will first discuss settings that affect availability and accuracy of geolocation data independent of operating system, then discuss the operating system-specific settings under various developers’ decisions and finally the user-defined settings. Choices made during development of the Cloudy with a Chance of Pain app are listed in the Methods section below.

Both operating systems can employ the GPS, network signals (a combination of WiFi and cell-tower signals, and, in some cases, signals of so-called Bluetooth beacons), or both to determine the smartphone user’s geolocation. These strategies
vary in accuracy, availability and battery power consumption. GPS provides the highest accuracy, but only works outside of buildings and uses high battery power. Deriving a user’s location based on network signals usually has a lower accuracy, but consumes less battery power and is available both inside and outside. Battery consumption also depends on frequency of location requests, the requested accuracy and the frequency of uploading location data to the application’s servers. App developer guides for both Android and iOS applications encourage app developers to choose settings that preserve battery power.

In both operating systems, apps can be ‘active’ (in use on the smartphone’s screen), ‘in the background’ (not in use, but not terminated), or ‘not running’ (not launched or terminated by system/user).

### Location services in Android

For Android, Google’s mobile operating system, app developers specify the logic for receiving location updates: what strategy to use (GPS, network or both), frequency, accuracy and the thresholds for requesting updates (time interval or change in distance). The app collects multiple location updates from the GPS and/or Network Location Provider for a specified amount of time, and then chooses the best estimate. The logic for best estimate is determined by the app developer, depending on the purpose of the app (e.g. most recent, most accurate). Usage of the GPS provider and size of window for collecting location updates are associated with highest power consumption.

In Android, location updates are available when the app is running, when the app is in the background, but not when the app is terminated. If a user ‘restricts background data’, location services (and other functionalities) are suppressed when a user is not connected to WiFi. Retrieval of location updates can automatically be resumed when the device is restarted [4].

### Location services in iOS

For iOS, Apple’s mobile operating system for iPhones, three strategies for determining user location exist: standard location service (GPS), the significant-change location service (network signals, but only supplied after a developer-specified ‘significant change’ in user’s location) and region monitoring (detects user’s entry and exit into specific regions). App developers specify the strategy, frequency of updates and thresholds for requesting updates. Compared to Android, iOS highly regulates geolocation requests when the app is in the background. Background processes are always time-restricted, making continuous data collection almost impossible. After an app is terminated, GPS geolocation is not available, but the significant-change location service can restart the app ‘in the background’ every fifteen minutes, reporting location changes as long as the operating system allows. If ‘Background App Refresh’ is disabled by the user, none of the location services are available when the app is not in use. Location updates can also be programmed by the developer to pause under certain conditions (e.g. when the user is unlikely to be moving) [5].

### User actions affecting location services

The smartphone user can also influence availability of geolocation data. On both operating systems, users can (a) disable location services for all apps in the general settings, (b) deny location services for the specific app, (c) switch off the device, (d) switch the device to a mode that does not allow location services (i.e. airplane mode, or, for Android phones, battery-saving mode), or (e) switch off WiFi/mobile network (only affects geolocation through network signals).

### Methods

#### Participants

UK residents who experienced pain for at least the preceding 3 months were invited to participate, through charity and patient organisations, and through study publicity on television, radio and in newspapers from January 20th 2016. Further requirements for participation were: Age 17 or over, resident in the UK and owning an Android (Android 3 or later) or iOS smartphone (iOS 8 or later). To enrol in the study, participants downloaded the uMotif Cloudy with a Chance of Pain app, consented to participate and completed a baseline questionnaire.

Ethical approval was obtained in December 2015 from the University of Manchester Research Ethics Committee 4 (ref: ethics/15522).

#### Cloudy with a Chance of Pain app

Every day, participants received a reminder to rate ten aspects of their symptoms in the app on a five-point ordinal scale. They could answer (any of) the ten aspects multiple times a day, for example in case of changing pain. In parallel, the smartphone’s GPS passively recorded geolocation up to hourly, thus ideally recording 24 geolocation observations per day.

A participant-day was defined as a day on which the application collected any data (i.e. symptom rating submitted by the participant, passively recorded geolocation, or both). We distinguished between participant-days with and without symptom data. Participant-days eligible for the main analysis were those on which a participant rated at least one of the ten symptom aspects; days on which a participant did not record any self-reported data, but the app did collect geolocation, were excluded.

The Android and iPhone app were developed to be as similar as possible in their geolocation retrieval strategy. The app uses both GPS (outdoors) and network signals (inside buildings). Every hour on the hour, it retrieves the last observed geolocation in the previous 60 minutes and its accuracy. If only network signals are available, the location update is only triggered in case of significant change (i.e. if a user is inside a building and does not move for an hour, the event is not fired). Geolocations and their timings are saved by the app until participants have internet connection, and then sent to the servers.

#### Analysis

Participants eligible for this analysis were those recruited between 20 January and 16 November 2016 and who had provided symptom data for at least 7 separate days. We analysed data completeness for all participant-days that met the inclusion criteria. We compared data completeness between Android and iOS participant-days, and the demographic characteristics between the Android and iOS users. The Mann–Whitney U test was used to test whether the number of geolocations per participant-day differed between operating systems and whether the number of symptoms per participant-day differed between operating systems.
Results

Subjects

Figure 1 shows the flowchart of study participants. A total of 12,441 people downloaded the smartphone application. We excluded participants who did not provide any symptom data at all \((N = 1222)\), that provided symptom data for less than 7 days \((N = 5383)\) or for whom we could infer that they manually disabled location services for our app, never allowing geolocation requests \((N = 205)\). Of the 5601 remaining participants, 47% used an Android device, 17% used an iOS device and for 37% the device type could not be determined because the users enrolled in the study before reporting of operating-system type was added to the app’s functionalities.

![Flowchart of number of participants.](image)

Figure 1 – Flowchart of number of participants. We excluded participants who never submitted symptom data, symptom data on less than 7 days or that did not give allow geolocation requests.

Participant-days

Figure 2 shows the participant-days. The 5601 included participants provided data on 492,080 days. We excluded 142,787 participant-days on which geolocation was collected, but participants did not provide symptom data. Of the remaining 349,293 days, 28% were provided by the iOS users, 46% were provided by the Android users and 26% were provided by users for which operating system was unknown. On average, participants submitted symptom data on 62 days. The median number of days was 33 (IQR: 14–76).

![Participant-days of the 5601 included participants.](image)

Figure 2 – Participant-days of the 5601 included participants.

Overall completeness of geolocation data

Figure 3 shows the distribution of completeness of geolocation data for days on which users submitted symptom data. Of all the participant-days, 16% had no geolocation data. Of the remainder, the most common number of geolocations was 24 (17%), followed by 1 (14%), then 2 (13%). Figure 4 shows the distribution of geolocation data completeness by known device type for all days of participants, as percentage of participant-days per device type. Most of the 101,714 participant-days collected by iOS devices collected either two geolocations (27%) or one geolocation (26%). For the 163,983 participant-days collected by Android devices, the most common were 24 geolocations (32%), followed by zero geolocations (18%). Table 1 shows that the median number of GPS observations was 17 for Android participant-days and 2.8 for iOS participant-days. The Mann–Whitney \(U\) test showed a statistically significant difference in distribution of number of geolocations per participant-day

![Geolocation data completeness by operating system.](image)

Figure 3 – Distribution of number of geolocations retrieved on 349,293 participant-days with symptom data.

![Distribution of number of geolocations, retrieved on 265,697 participant-days for Android users (blue bars; \(N = 101,714\) participant-days) and 972 iOS users (red bars; \(N = 163,983\) participant-days) that did not disable location services over the complete study period. Participant-days with an unknown operating system are excluded.](image)

Figure 4 – Distribution of number of geolocations, retrieved on 265,697 participant-days for Android users (blue bars; \(N = 101,714\) participant-days) and 972 iOS users (red bars; \(N = 163,983\) participant-days) that did not disable location services over the complete study period. Participant-days with an unknown operating system are excluded.
between both operating systems ($p < 0.01$). However, the difference in number of symptoms submitted by the participants (mean 10.4 versus 10.3) is not practically meaningful. This difference suggests that an average Android user submits one complete symptom rating more than an iOS user after 100 days of symptom data entry. The statistically significant difference in symptom ratings between the two operating systems can be explained simply by the large volume of data.

Table 1 – Summary of data completeness of geolocation and self-reported data completeness per participant-day for Android and iOS users. Number of symptom data entries per participant-day refers to the number of symptom aspects rated by that participant (one complete submission consists of 10 symptom aspects).

<table>
<thead>
<tr>
<th></th>
<th>Android $N = 163,983$</th>
<th>iOS $N = 101,714$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geolocation</td>
<td>Mean 14.1 2.8</td>
<td>Median 17 2</td>
</tr>
<tr>
<td>Symptoms</td>
<td>Mean 10.4 10.3</td>
<td>Median 10 10</td>
</tr>
</tbody>
</table>

**Discussion**

We investigated data completeness of a large mobile Health (mHealth) study in which participants rated ten aspects of their chronic pain in a smartphone app that simultaneously collected their geolocation. Although the app was developed to retrieve 24 geolocations per day (i.e. hourly), it was only successful on 17.3% of the participant-days. Part of the missing data can be explained by users switching to airplane mode or battery-saving mode during the day. Users may also have manually terminated the app, in which case geolocation updates are only triggered in case of ‘significant change’. Participants are unlikely to make these significant location changes when they are sleeping, and, when they are at work for those with sedentary jobs. Another possible explanation is the timing of the prompt to submit symptoms – 18:24 in the evening – which was expected to fit best into participants’ lifestyles. If participants manually terminate the app and then stay put for the evening, the location strategies not requiring ‘significant change’ would only be launched after 18:24 in the evening, resulting in missing data in the hours before. Especially interesting is the high prevalence of days with no geolocation updates. Following the app-developers’ settings, we would expect at least one geolocation to be retrieved when the participant submitted symptom data, even if the device had not ‘significantly changed location’ before that time. As many people charge their phones overnight, battery life is expected to decrease over the day, increasing the chance that the phone is in battery-saving mode at 18:24.

Distribution of data completeness differed greatly between the two operating systems. Specifically, an average of 14.1 geolocations was collected on participant-days of Android users, compared to 2.8 for iOS participant-days. As there was no meaningful difference in symptom-entry behaviour between Android and iOS participants, this difference is probably due to the tendency of iOS to block data processing when an app is ‘in the background’. Although percentages of iOS participant-days with few (1–9) geolocation retrievals were higher, a higher percentage of Android participant-days had 0 geolocation retrievals. This may be because historically, only Android phones had a battery-saving mode. This functionality was only introduced in iOS version 9, briefly before the start of this study (16 September 2015), but cannot be installed on the iPhone 4 and older models.

A limitation of this study was that we had to exclude 2047 participants whose operating system was unknown because they joined the study before the functionality of recognising device type was added to the app. Furthermore, we could not determine the cause of missing geolocation data because our application does not collect information on app status or any user-defined settings.

We have not found mHealth studies that are comparable to Cloudy with a Chance of Pain in number of participants, operating system, and geolocation sampling frequency. The most similar study was a feasibility study of using smartphones to investigate air-pollution exposure [7]. This study had a smaller sample size ($N = 54$) and only developed an application for Android phones. They did not use Android’s positioning system, but used external software called Skyhook to collect geolocation more frequently (every 5 minutes). Their study lasted for three months. Due to differences in reporting, it is hard to compare our study with theirs. For example, they excluded 10 participants that did not keep manual location diaries and 4 participants who did not report having their phone powered on 24 hours a day from the analysis. Of the remaining 38 participants, 4 participants had, on average, complete geolocation data (mean number of geolocations per day 288, every 5 minutes) and another 26 participants had an average of 144 or more (i.e. geolocation every 10 minutes). Only one participant had a daily average under 72 data points, probably still providing more granular data than participants in our study. The authors found that geolocation data completeness depended on smartphone manufacturer, mobile network and the time of day data points were collected, as well as whether the user switched the device off or to airplane mode.

Although data completeness of smartphone-collected geolocation may be suboptimal in our study, various strategies may enable mHealth researchers to still perform the intended investigation. Investigating the richer Android geolocation data to infer information on the sparser iOS data might be possible. By analysing individual patterns in geolocation, recurring periods in which no significant change in location takes place may be identified.

The results of as [7] indicate, however, that even when developing an app for Android users only, data completeness may not be optimal. As iOS’s market share is 44% in the UK (large, compared to Brazil’s 4% and China’s 17% [6]), including both Android users and iOS users in medical studies, may reduce selection bias. On the other hand, any systematic differences between iOS users with sparse geolocation data and Android users with richer geolocation data could also compromise the study’s validity.

In conclusion, this study shows that mHealth researchers collecting geolocation via participants’ own smartphones may still encounter missing data, especially in participants that use iPhones rather than Android devices. As future work we plan to investigate robust methods for imputing geolocation data. Researchers should develop research smartphone applications with the operating systems’ specifications in mind and instruct participants to refrain from configuring their phone to settings that do not allow geolocation requests. In addition, the mHealth research community should devise standards for reporting geolocation data quality and analysing systematic differences between participant groups.
Conclusion

We have shown that a smartphone app, when sampling the user’s geolocation at regular intervals, does not always collect hourly geolocation. Missing data is more common in iPhone users compared to Android users.

To increase the data quality of geolocation data, mHealth researchers should be conscious that geolocation availability and accuracy depends upon the operating system, app developer and user. Smartphone applications should be designed with these factors in mind.

Our results indicate that geolocation sampling may also be suppressed by the operating system, especially on iOS devices, leading to missing data outside of researchers’ control. Methodologies to assess – and possibly correct for – systematic differences in data quality between groups of participants should, therefore, be devised.

Acknowledgements

The study was supported by Arthritis Research UK (grant reference 21225), the Arthritis Research UK Centre for Epidemiology (grant reference 20380) and the Farr Institute @HeRC (MR/K006665/1). This report includes independent research supported by the National Institute for Health Research Biomedical Research Unit Funding Scheme. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

References


Address for correspondence

Anna Beukenhorst, Room 2.811, Stopford Building, Faculty of Biology, Medicine and Health, Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Institute of Inflammation and Repair, The University of Manchester, Oxford Road, Manchester, M13 9PT