

Annex 3

Ofcom mobile research application: technical and statistical methodologies

- 1.1 The Ofcom Mobile Research app project is the next phase of Ofcom's work to measure mobile performance and consumer experience of using mobile services. The new methodology established a panel of UK smartphone users who installed an Ofcom-branded research app on their Android smartphone. The Ofcom Mobile Research app was developed by P3 communications GmbH, Ofcom's technical partner in this research. P3 also managed the collection and aggregation of the collected data, and helped with data analysis.

About the app

- 1.2 The app is available on Google's Play Store. When opening the Research app for the first time, a pop-up takes users into their smartphone setting, where they need to add the Ofcom Research App to the list of 'apps with usage access'. This additional type of permission is required to ensure proper and reliable working of the Research App. The app then asks users to consent to the data collection and to provide demographics information.
- 1.3 The app collects data passively, as people use their phones, and does not require user interaction. A subset of the collected data is presented in an aggregate view to the user within the app's user interface (shown below).



- 1.4 An important part of the research is to compare mobile network performance with customer satisfaction levels, therefore, the app from time to time presents pop-ups:
 - 1.4.1 After installing the app, a 'Welcome' pop-up is displayed, saying that the app will present pop-ups from time to time; those are described in (2) to (5) below.

- 1.4.2 At the beginning of the research and then every 30 days, the app asks for overall satisfaction with mobile service provider.
- 1.4.3 At the beginning of the research, the app asks to rate the importance of specific services (such as voice, download/upload, browsing and audio/video streaming).
- 1.4.4 From time to time (based on predefined trigger frequencies), the app also asks users to rate their satisfaction with voice calls and apps used. Satisfaction survey pop-ups are triggered only for the following applications:

Phone call	WhatsApp	YouTube	eBay
Facebook	Facebook Messenger	Play Store	Gmail
Twitter	Viber	Chrome	Stock Android mail
Instagram			

- 1.4.5 On day 15 of the research and then every 30 days, a pop-up is displayed, reminding that the app is still installed on the phone.
- 1.4.6 The prompts for (2) and (3) appear in the notification tray. Answering a pop-up survey (4) is as simple as clicking one button with satisfaction corresponding to one of the three emoticons shown.

Data collection

- 1.5 P3’s data collection framework is designed to only collect anonymous information. Any collected data is stored within the app’s sandbox while resting on the device and using encrypted transmission when offloading the information to the backend systems. The app has successfully passed an external security audit commissioned by Ofcom.
- 1.6 One major objective of this research method was to have minimal impact on the consumer’s day-to-day smartphone use – particularly with regard to battery consumption and mobile data usage. The app will use a maximum of 3MB of data per month for its active connection test, and will only send the data collected over a mobile network if it cannot send it over Wi-Fi. The corresponding average data used depends on users’ overall smartphone use, but is typically no more than 1MB per day across mobile network and Wi-Fi. The actual background data collection is also designed to have minimal interference with customer experience.
- 1.7 The data collection can be broadly grouped into the following use cases:
 - 1) Measurement results
 - Data service availability and accessibility
 - Voice calls
 - Data throughput and responsiveness (latency)
 - Behavioural and usage information

- 2) User input (e.g. satisfaction surveys)
- 3) Technical parameters
 - Location information¹
 - Device and system parameters

1.8 Data collection runs 24/7 and should only be interrupted by phases where the device is switched off. Most data points are collected whenever the device is actively used, e.g. whenever the user is interacting with apps or making phone calls. In addition, an automated “light-weight” connection test is attempted four times an hour to test the data service availability.

Calls dropped due to loss of service

- 1.9 The passive data collection includes information on the voice service for calls made or received by a panelist. The framework is triggered on the use of the legacy telephony manager (app). More precisely, when the user attempts to setup a call by pressing the green calling button or when accepting an incoming call. Data is collected every second until the call is ended. The voice call types taken into consideration include legacy circuit-switched calls (2G, 3G) as well as cases of circuit-switched fallback (calls initiated while the device was still on 4G and then immediately handed off to 3G or 2G) and VoLTE (calls made over 4G networks).
- 1.10 The second-by-second data logs contain, amongst other things, radio information, such as cell ID, radio network technology and signal strength, as well as information on call duration and call-end-type.
- 1.11 In controlled measurement setups like drive tests, one of the most common voice metrics is “dropped call ratio”. The ‘over-the-top’ crowdsourcing approach chosen in this research has to accept limitations implied by the availability and accessibility of information exposed by the smartphone operating system. Due to security precautions taken by the operating system developers and device manufacturers, not all registers can be accessed and read from regular commercial devices by third party software.² One of those non-accessible registers is the call-end-type and its exact technical release cause.
- 1.12 P3 has developed a method to detect cases where the call is interrupted due to a loss of service, however, such an event can only be detected for the party that had the Ofcom research app installed. This metric hence establishes a subset of the dropped call ratio used in active testing.

Data network share

- 1.13 The passive data collection includes information on the actual apps being used. The framework is triggered when an app is actively pulled into foreground, i.e. opened

¹ To ensure battery is not affected, the app mostly relies on the last known location available from the Android OS. This might be a GPS-grade position, whenever another app has used GPS. The only time the app actively requests a GPS-grade position is when a voice call is made or received, to be able to locate the exact position of any dropped calls.

² Technically, a device can be tweaked to expose more registers, e.g. if it gets “rooted” and the app assigned with system privileges. However, this approach cannot be accommodated in the specific crowdsourcing approach chosen.

and used by the user. Information is collected per second while an app is in foreground and transferring data (sending and/or receiving).

- 1.14 The second-by-second logs contain, amongst other things, radio and traffic information, such as cell ID, radio network technology, signal strength and the transferred data volume. The radio network technology is tagged by the wireless network technology, i.e. 2G, 3G, 4G or Wi-Fi.
- 1.15 The data network share metric establishes the ratio of seconds per network technology across all used apps, and hence expresses the time share on 2G, 3G, 4G and Wi-Fi while actively using apps and transferring data.

Statistical methodology

- 1.16 This report analyses findings collected from panellists who had the Ofcom Mobile Research app downloaded for at least 7 days during the first fieldwork period. Panellists were included if they had valid demographic details, such as age, gender and location, are users of the EE, O2, Three and Vodafone networks and have access to 3G and/or 4G mobile technologies. The overall panel who match these criteria comprised of 4,288 people. The panellists were distributed as follows:
 - By Mobile Network Operator
 - 1,385 users of EE
 - 1,005 users of Three
 - 857 users of Vodafone
 - 1,041 users of O2
 - By Nation
 - 3,440 panellists with a home postcode in England
 - 226 panellists in Northern Ireland
 - 350 panellists in Wales
 - 272 panellists in Scotland
 - By Rurality
 - 787 panellists had a home postcode in a rural area
 - 3,501 panellists had a home postcode in an urban area
- 1.17 The first fieldwork reporting period for this research includes data collected from 27th September until 23rd December 2016. During this time, over 40 million rows of anonymised data were collected passively with very little interaction needed.
- 1.18 Data are collected on data service availability and performance, reliability of voice calls, customer experience and customer satisfaction.

Sampling methodology

- 1.19 This is Ofcom's first mobile research app downloadable from Google Play Store. The research app is advertised on the Ofcom website, in the Technology press and through social media. Although the app is crowd-sourced, the research project is panel based with sampling targets for each of the four Mobile Network Operators (EE, Three, O2 and Vodafone) and for different demographic and geographic groups.

Reporting panels

- 1.20 At the inception of the research it was decided that Ofcom would not report findings on the whole panel, but that it would report on three sub-panels that selected panellists to be representative by Mobile Network Operator (MNO), demography and geography. The two sub-panels' findings in this report are from the:
- Provider Performance Panel
 - UK Performance Panel
- 1.21 There is also a Behavioural sub panel with 1,200 panellists selected to be representative by demographics and MNO market share. The results for this panel will be used to look at how consumers use their mobile phone. This data will be analysed and reported on in subsequent reports.
- 1.22 Both the Provider Performance panel and UK Performance panel are defined below. The location for membership to each panel is defined by home postcode, however, any analysis carried out is based on the location of where readings are taken.

Provider Performance Panel

- 1.23 This panel was designed to allow comparisons of consumer experience across mobile providers. The panel has an equal number of panellists per MNO (650 panellists), who are deemed to have access to 4G technology, and then representative³ by UK nation and rurality. If a panellist has any readings where 4G technology was accessed, then that panellist is deemed to be a '4G panellist'.
- 1.24 Readings for 2G, 3G and 4G mobile technologies (and Wi-Fi) are analysed.

Provider panel sample proportions and numbers

MNO: EE: 650 Three: 650 Vodafone: 650 O2: 650

Of which -

Nation: Scotland: 8% Wales: 6% Northern Ireland: 3% England: 83%

Rurality: Rural: 14% Urban: 86%

UK Performance Panel

- 1.25 The purpose of the UK Performance Panel is to monitor consumer experience at a UK level. The panel was designed to be representative by MNO market share, UK nation and UK rural/urban location. As with the Provider Performance panel, 2G, 3G and 4G readings (and Wi-Fi) are analysed.

³ ComScore MobiLens Plus Audience Profile May 2016 data for Android users was used to estimate geographic representativeness across MNOs. UK census 2011 data was used to estimate rural and urban percentages across all MNOs.

- 1.26 The 4G UK Performance Panel (where panellists are deemed as having access to 4G technology) comprises of 2,904 panellists. The panel is representative by MNO market share and geography. Data is analysed at the overall UK level, by rurality and/or by nation.

UK Performance panel sample proportions

MNO: EE: 39% Three: 16% Vodafone: 21% O2: 25%

Nation: Scotland: 8% Wales: 6% Northern Ireland: 3% England: 83%

Rurality: Rural: 14% Urban: 86%

Metrics and analysis methodology

- 1.27 For *Comparing Service Quality*, three metrics were focused on:

- Calls dropped due to loss of service
- Data service availability
- Data network share

- 1.28 For details on how these metrics are calculated please see the Technical Methodology.

- 1.29 There were a very large number of readings collected during the fieldwork period. The number of panellists contributing readings to each metric was checked to make sure that the sample size was large enough as to be statistically robust. For all the metrics, the sample size for the smallest sub group analysed is at least 95.

Calls dropped due to loss of Service and Data service availability

Provider Performance Panel

- 1.30 The Provider Performance panel was used to compare: “Calls dropped due to loss of service” and “Data Service Availability” by MNO. It was decided that differences in proportions between MNOs would only be reported if they are significantly different as judged by a 5% test of statistical significance.
- 1.31 As the analysis is carried out directly on the number of readings, and not the number of panellists, comparisons were carried out on hundreds of thousands of readings. Because of the very large number of readings, it is essential to not just look at whether the results are statistically significantly different but also to look at the size of the difference.
- 1.32 When comparing hundreds of thousands of readings, even very small differences between averages or proportions can be found to be statistically significant by running a statistical test. However, this may not actually be a noticeable or practical difference for consumers’ experience of using their phones unless the size of the difference is large enough. When running a statistical test, the effect size shows the size of the difference between groups.

- 1.33 For this research, it was decided not to report on differences unless the findings are both statistically significant at the 99% level of confidence and that there is at least a medium effect size (0.3). A large effect size has a value of at least 0.5.
- 1.34 In our analysis, although the comparisons are statistically significantly different by MNO, the effect size is too small to report a practical difference (less than 0.1 for both metrics). We conclude that the data shows very little difference between MNOs and therefore data by MNO is not presented.

UK Performance Panel

- 1.35 The UK Performance Panel was used to look at the metrics at the overall UK level, by rurality and by nation. The same methodology for looking at sample sizes and deciding on significant differences was used as for the Provider Performance panel.
- 1.36 There are only very low effect sizes (less than 0.1) for both metrics when comparing readings for rural vs urban areas and by nation. Therefore, data for both metrics are presented at the UK level only.

Data Network Share

- 1.37 Data Network Share is only presented in this report at an overall UK level as it is thought to also be driven by consumer behaviour as well as the performance of the mobile network. How often a panellist phone connects to each mobile technology and Wi-Fi can be due to the customisation of phone settings to detect and connect to Wi-Fi as well as the availability of the mobile data network. Therefore, no comparisons by MNO or location are presented.