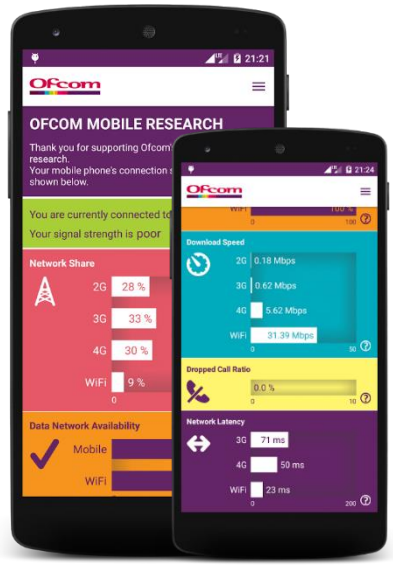
A1. Technical methodology

The Ofcom mobile research app project is the latest phase of Ofcom's work to measure mobile performance and the 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[1]. This 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

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 settings, where they are asked to add the Ofcom mobile research app to the list of 'apps with usage access'. This additional permission is required to ensure proper and reliable working of the app. The app then asks users to consent to data collection and to provide demographic information.

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] The operating system used on iPhones (iOS) has restrictions on apps running in the background and being able to access network performance data. Therefore, an iPhone version of the app is currently not available.

Part of the research is to analyse customer satisfaction levels, so from time to time the app presents pop-ups:

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.
2. At the beginning of the research and again every 30 days, the app asks for the user's overall satisfaction with the network performance of their mobile service provider.
3. At the beginning of the research, the app asks the user to rate the importance of specific services (such as voice call, download/upload, browsing and audio/video streaming).
4. From time to time (based on predefined trigger frequencies), the app asks users to rate their satisfaction with voice calls and apps used. Satisfaction survey pop-ups are triggered only for the following applications – one in every 38 except for WhatsApp (which is one in every 75):

Phone call	WhatsApp	YouTube	Chrome	Twitter
Facebook	Facebook Messenger	Play Store	Gmail	Instagram

5. On day 15 of the research and again every 30 days, a pop-up is displayed, reminding the user that the app is still installed on the phone.

The prompts for 2. and 3. above 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

P3's data collection framework is designed to collect only 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 back-end systems. The app has successfully passed an external security audit commissioned by Ofcom.

A major objective of this research method was that it had minimal impact on the consumer's day-to-day smartphone use – particularly with regard to battery consumption and mobile data use. The app uses a maximum of 3MB of data per month for its active connection test, and only sends the data collected over a mobile network if it cannot send it over wifi. The corresponding average data used depends on users' overall smartphone use but is typically no more than 1MB per day across mobile network and wifi. The background data collection is also designed to have minimal impact on the user's experience.

The data collection can be broadly grouped into the following use cases:

1. Measurement results
    - Data service availability and accessibility
    - Voice calls
    - Data speed and responsiveness (latency)
    - Behavioural and usage information
2. User input (e.g. satisfaction surveys)
3. Technical parameters
    - Location information[2]
    - Device and system parameters

Data collection runs 24/7 and is interrupted only by phases when the device is switched off. Most data points are collected whenever the device is in active use; e.g. whenever the user is interacting with apps or making phone calls. In addition, an automated 'lightweight' connection test is attempted four times an hour to test the data service availability.

Metrics

Calls dropped due to loss of service

The passive data collection includes information on the voice service for calls made or received by a panellist. The framework is triggered on the use of the legacy telephony manager (app); more precisely, when the user attempts to set up a call by pressing the green calling button or when they accept an incoming call. Data are collected every second until the call ends. The voice call types taken into consideration include legacy circuit-switched calls (2G, 3G), 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).

The second-by-second data logs contain, among 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.

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 for this research must accept the 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[3]. One of those non-accessible registers is the call end type, and its exact technical release cause.

[2] To ensure the 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.
[3] Technically, a device can be tweaked to expose more registers, e.g. if it gets 'rooted' and the app is assigned with system privileges. However, this approach cannot be accommodated in the specific crowdsourcing approach chosen.

P3 has developed a method for detecting cases where the call is interrupted due to a loss of service, although such an event can be detected only for the mobile user who has the Ofcom research app installed. This metric therefore establishes a subset of the dropped call ratio used in active testing.

Data service availability

The background data collection includes an automated, active test case, also referred to as the 'connection test'. This connection test is a lightweight measurement, testing the availability and accessibility of a data connection to a cloud-based server in the internet. The test is attempted every 15 minutes, regardless of the available wireless network technology at the time of the test.

The test attempts to download a 596 bytes-sized icon via a static URL, hosted on Amazon's CloudFront, a global network service for hosting content (also known as content delivery network, or CDN).

Three phases need to be completed for a test to be considered successful:

a) the resolution of the URL via DNS[4], to obtain the IP address of the server from which the test will attempt the download;

b) the TCP[5] connection setup; and

c) the actual download of the content.

Each phase is logged with its respective success or failure. The evaluation of data service availability considers the success or failure on the final phase, i.e. the attempt to download the icon from the destination server.

For the connection test to be executed, a couple of prerequisites need to be fulfilled. In the light of the conservative approach to battery consumption, the execution of the test is not attempted if the device battery level is at 15% or less. Once the battery is recharged to a level above 15%, the test automatically resumes without user interaction. Different versions of operating systems and respective default and/or user-enforced settings also need to be taken into account. This may result in a lower number of tests being executed, or in tests needing to be removed from the analysis.

The measurement data collected include details on the device, the operating system and other relevant system and environmental parameters, e.g. radio information that allows us to separate cellular from wifi tests.

As a result, the data service availability metric, established for this report, provides a view on the success or failure to connect to the internet using a 2G, 3G or 4G mobile data network connectivity.

[4] DNS stands for domain name system. It is a worldwide, distributed directory service and is used in this context to resolve and translate the domain name of the destination URL (xyz.cloudfront.net) to the numerical IP address required for locating and identifying the destination server with the underlying network protocols.

[5] TCP stands for transmission control protocol. It is one of the main protocols of the Internet protocol suite. The TCP connection phase is mainly characterised by a multi-step handshake process between the client and the server (connection establishment) before entering the data transfer phase.

Response time

Response time (referred to technically as latency) is the delay between a consumer making a request to their mobile network for information and the network providing this information to the device.

Latency is measured as part of the connection test (described in detail in the section above). Once a successful resolution of the server IP address is logged (DNS lookup – first phase of the test), a flight of ten ICMP pings (internet control message protocol) is sent towards the same destination server as for the connection test (Amazon's CloudFront).

The following ping command is used to send ten pings with 200ms pause in between, and an overall timeout of 30 seconds:

ping –i 0.2 –c 10 –W 30 [… server ip …]

The ping response is parsed for obtaining the round-trip times.

Data performance: passive speed test

As with previous metrics, data performance is measured passively. For this metric, we evaluate the actual data amount transferred while panellists are using apps as part of their everyday phone use.

Traffic information is logged for each second when the app in foreground (i.e. the app a person is using) is transferring data (uploading or downloading). P3's patent-pending approach, run as part of the backend processing on the collected stats, calculates the so-called passive speed test samples (PSTs) from the foreground app traffic bursts.

The PSTs describe a customer-centric view of the speeds as demanded by the apps used and as defined by the user's setup – mainly device and tariff. It is important to note that this speed does not reflect network capabilities but instead looks at the actual user experience. The speed recorded is thus lower than that from network performance measurements or speed tests, which look at the best service a network can offer.

Data network share

The passive data collection includes information on the actual apps being used. The framework is triggered when an app is actively pulled into the foreground, i.e. opened and used by the user. Information is collected per second while an app is in the foreground and transferring data (sending and/or receiving).

The second-by-second logs contain, among 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 wifi.

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 wifi while actively using apps and transferring data.