



Google Data Collection on Android

Darryn Campbell

SW Architect, Zebra Technologies
@darryncampbell

July 3rd 2019

Google Data Collection on Android

Who am I?

- Developer advocate / Software architect for Zebra Technologies
 - Android OEM developing task specific devices
- Feature owner for “GMS Restricted”



Google Data Collection on Android

2 Schools of Thought...

Support The Guardian

Available for everyone, funded by readers

Contribute →

Subscribe →

Search jobs

Dating

Sign in

Search ▾

The Guardian

UK edition ▾

News

Opinion

Sport

Culture

Lifestyle

More ▾

The Guardian view Columnists Cartoons Opinion videos Letters

Opinion

Facebook

Fri 30 Mar 2018 08.17 BST

59,796

3,024

Are you ready? Here is all the data Facebook and Google have on you

Dylan Curran

The harvesting of our personal details goes far beyond what many of us could imagine. So I braced myself and had a look

Editorially independent, open to everyone

We chose a different approach – will you support it?

Support The Guardian →

most viewed

Labour expels Alastair Campbell from party

Live Calls for Tory Islamophobia inquiry as equality body launches Labour antisemitism investigation - live news

Google Data Collection on Android

2 Schools of Thought...

Google knows where you've been

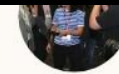
Google stores your location (if you have location tracking turned on) every time you turn on your phone. You can see a timeline of where you've been from the very first day you started using [Google](#) on your phone.

Click on this link to see your own data: google.com/maps/timeline?...

Here is every place I have been in the last 12 months in Ireland. You can see the time of day that I was in the location and how long it took me to get to that location from my previous one.



▲ 'A Google map of every place I've been in Ireland this year.' Photograph: Dylan Curran



eliminated in the EU
elections. Here's how we
helped do it
Nick Lowles

Google Data Collection on Android

2 Schools of Thought...

[Guardian account](#) to join the discussion.

Due to the large number of comments, they are being shown 100 per page.



UnstoppableSteve 28 Mar 2018 11:11

192 ↑

These things do appear to be configurable, though.

I've just checked a bunch of these repositories for my Google account and they're completely empty, even though I'm a regular user of youtube, Google Maps, Google, Gmail etc.

Google does appear to respect your configuration choices to opt out of personalised ads, location history etc

[Share](#)

Report



tomper2 → **UnstoppableSteve** 28 Mar 2018 11:31

153 ↑

Do you really trust Google?

[Share](#)

Report



justanotherflyboy → **tomper2** 28 Mar 2018 11:56

75 ↑

nope!

[Share](#)

Report



JosephBloggworthy → **UnstoppableSteve** 28 Mar 2018 12:06

102 ↑

Do you imagine those settings actually stop them from collecting the data, or just tell them not to use it to personalise the ads they show you?

Facebook and Google are known to maintain profiles on people that don't even have accounts with them by hoovering up snippets of information from other people's interactions with them. They'll never willingly forego an opportunity to add to their data trove. If you tick the box that says "don't track my location", they'll still track your location, they just won't use that information to suggest shops or restaurants, and won't add it to the location history they show to you when you ask.

The big tech companies have a golden rule that they share with the NSA - "CAPTURE EVERYTHING"

[Share](#)

Report

Google Data Collection on Android

Performing your own tests

Deciding on a Test Methodology:

Google Data Collection on Android

Performing your own tests

Deciding on a Test Methodology:

- What version of Android?
- What device hardware?
- Enable Location Services?
- High accuracy vs Battery Saving location?
- Perform Google searches?
- Use the Google Assistant?

Google Data Collection on Android

Performing your own tests

Deciding on a Test Methodology:

- What version of Android?
- What device hardware?
- Enable Location Services?
- High accuracy vs Battery Saving location?
- Perform Google searches?
- Use the Google Assistant?
- Use Google Now?
- SUW: Agree to share diagnostics?
- Sign in with a Google Account?
- Managed Android device?
- Multiple users?
- Should my test subjects move?

Google Data Collection on Android

Performing your own tests

Deciding on a Test Methodology:

- What version of Android?
- What device hardware?
- Enable Location Services?
- High accuracy vs Battery Saving location?
- Perform Google searches?
- Use the Google Assistant?
- Use Google Now?
- SUW: Agree to share diagnostics?
- Sign in with a Google Account?
- Managed Android device?
- Multiple users?
- Should my test subjects move?
- Should I move the device?
- Run / Walk / Cycle with the device?
- Should I actively use the device?
- Use a device with a SIM?
- Enable Bluetooth?
- Enable Chrome Diagnostics?
- Sign-in to Chrome?
- Enable Keyboard diagnostics?
- Enable apps to access location?

Google Data Collection on Android

Performing your own tests

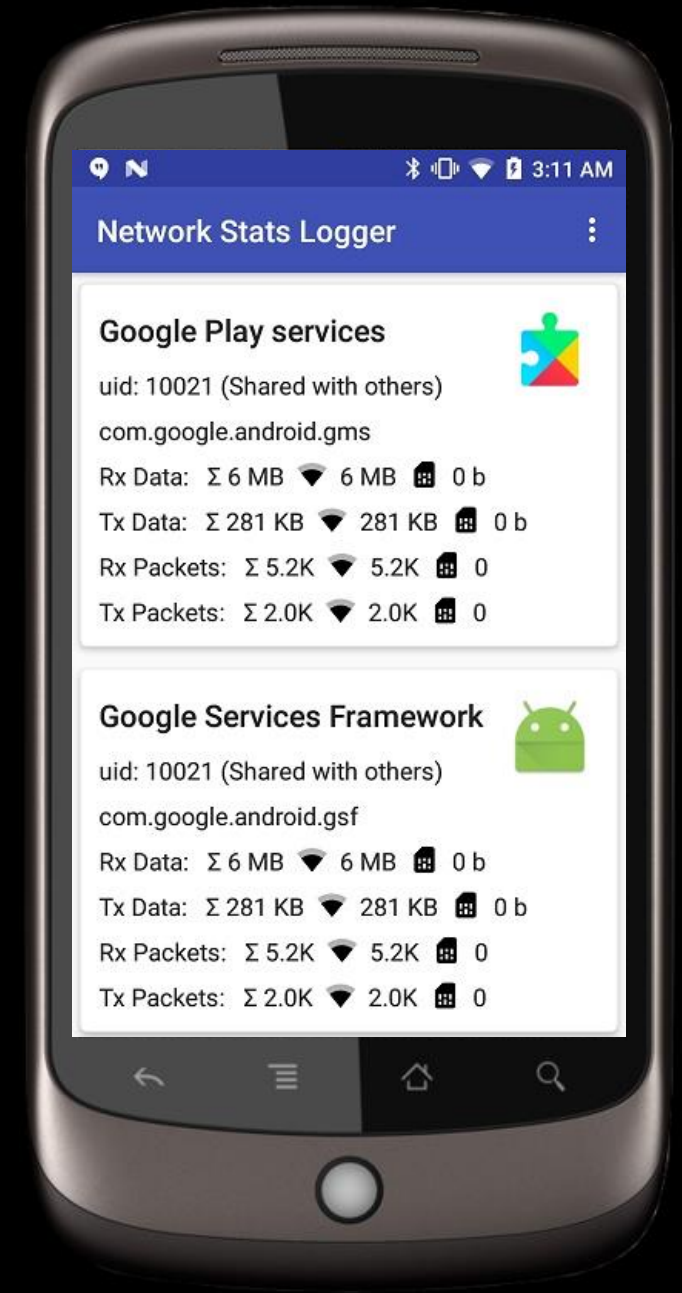
Deciding on a Test Methodology:

- What version of Android?
- What device hardware?
- Enable Location Services?
- High accuracy vs Battery Saving location?
- Perform Google searches?
- Use the Google Assistant?
- Use Google Now?
- SUW: Agree to share diagnostics?
- Sign in with a Google Account?
- Managed Android device?
- Multiple users?
- Should my test subjects move?
- Should I move the device?
- Run / Walk / Cycle with the device?
- Should I actively use the device?
- Use a device with a SIM?
- Enable Bluetooth?
- Enable Chrome Diagnostics?
- Sign-in to Chrome?
- Enable Keyboard diagnostics?
- Enable apps to access location?
- Install 3rd party apps?
- Browse the web?
- Which websites??
- Check email with Gmail?
- Send SMS with the messaging app?
- Make an appointment with Google Calendar?
- Use Google Maps to navigate?
- Sign into Google Maps?
- Enable Location History?
- Enable Google Backup?
- Enable BLE / WiFi scanning?
- Go near BLE beacons?
- Use linked Google Home?
- Browse content on YouTube?
- Enable Play Protect?
- Disable Play Services?

Google Data Collection on Android

Performing your own tests

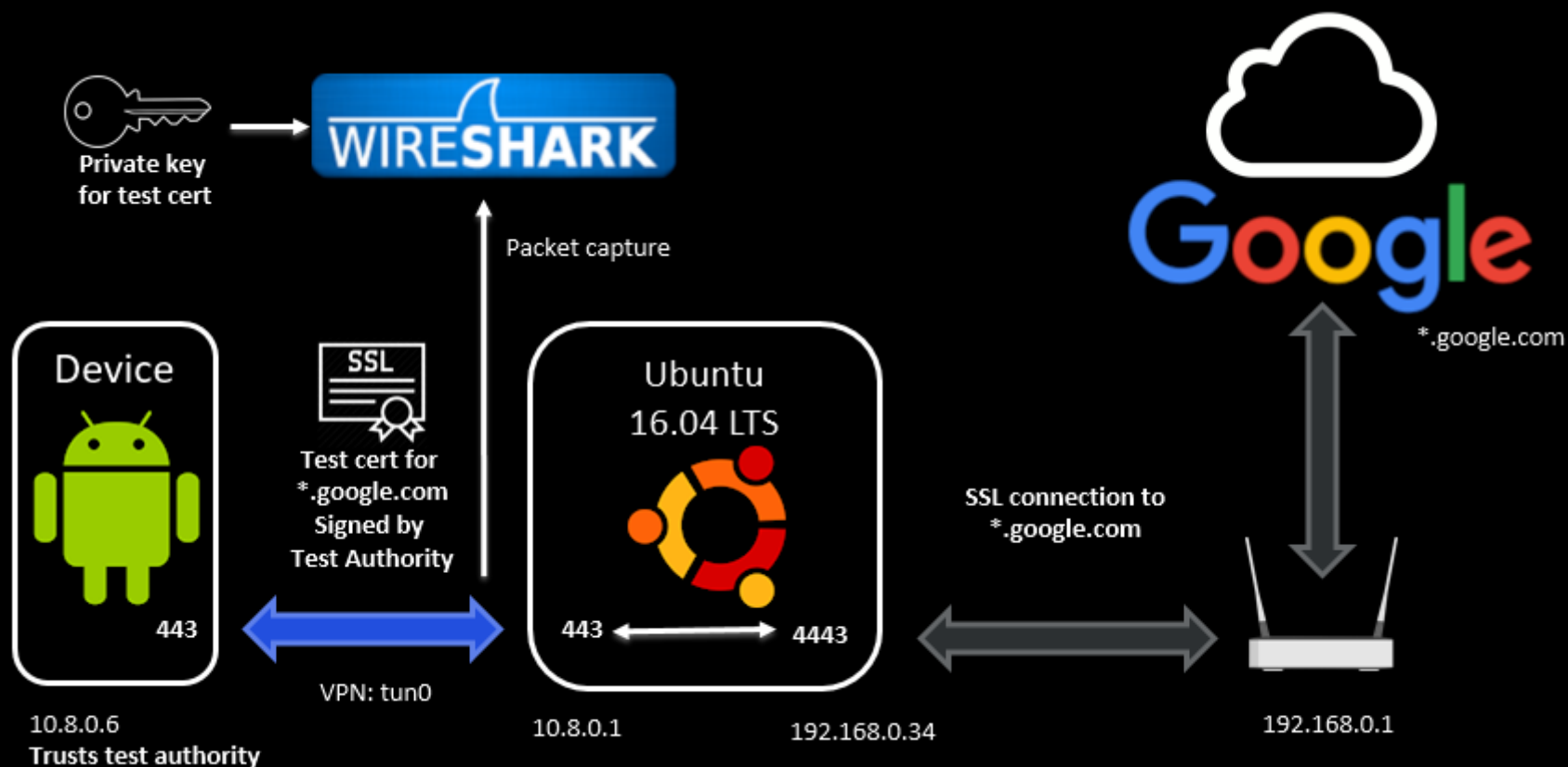
- Can tell how much data but not WHAT data is sent
- Quantitate Analysis:
 - Use the [Network Stats API](#)
 - Thousands of such apps – here's [one I wrote](#) ([Source code](#)).



Google Data Collection on Android

Performing your own tests

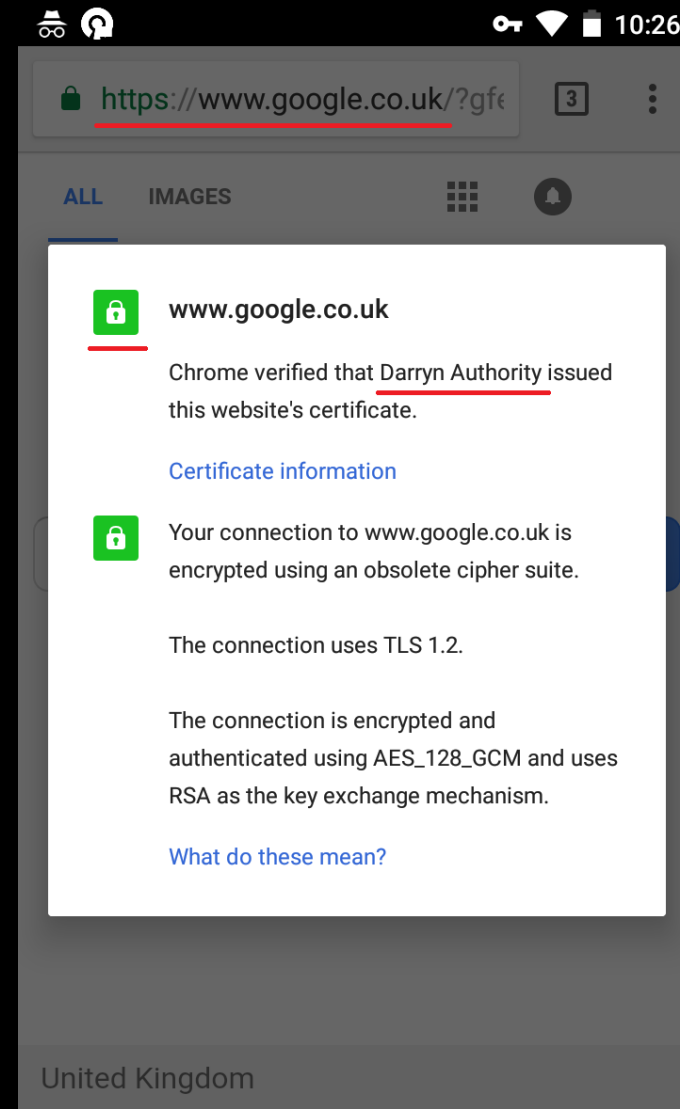
- Qualitative Analysis (What is being sent?):
 - <http://www.darryncampbell.co.uk/2017/12/29/android-traffic-analysis-to-google-servers-methodology/>



Google Data Collection on Android

Performing your own tests

- Qualitative Analysis:



Google Data Collection on Android

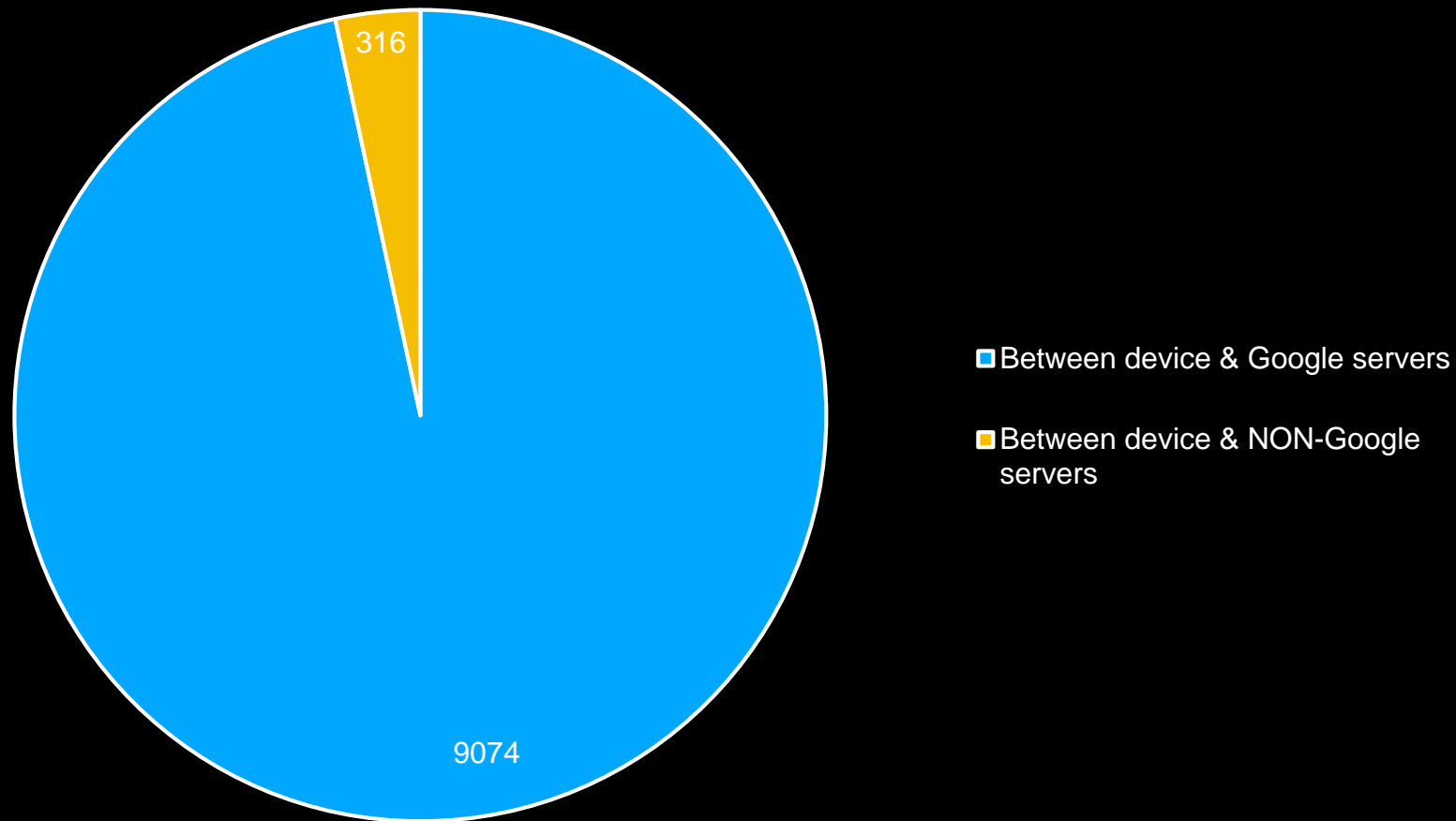
Some Data – My own analysis

- Test Conditions:
 - Android Nougat 7.1
 - Device kept:
 - Stationary
 - Untouched
 - Plugged into power
 - No applications manually launched
 - Google (@gmail) account exists on device
 - Factory Reset and run through SUW
- SUW Conditions:
 - Back up to Google Drive: Enabled
 - Help apps find location: Enabled
 - Improve location accuracy: Enabled
 - Send System Data (analytics): Enabled
 - Location History: Enabled
- Leave device overnight

Google Data Collection on Android

Some Data – My own analysis

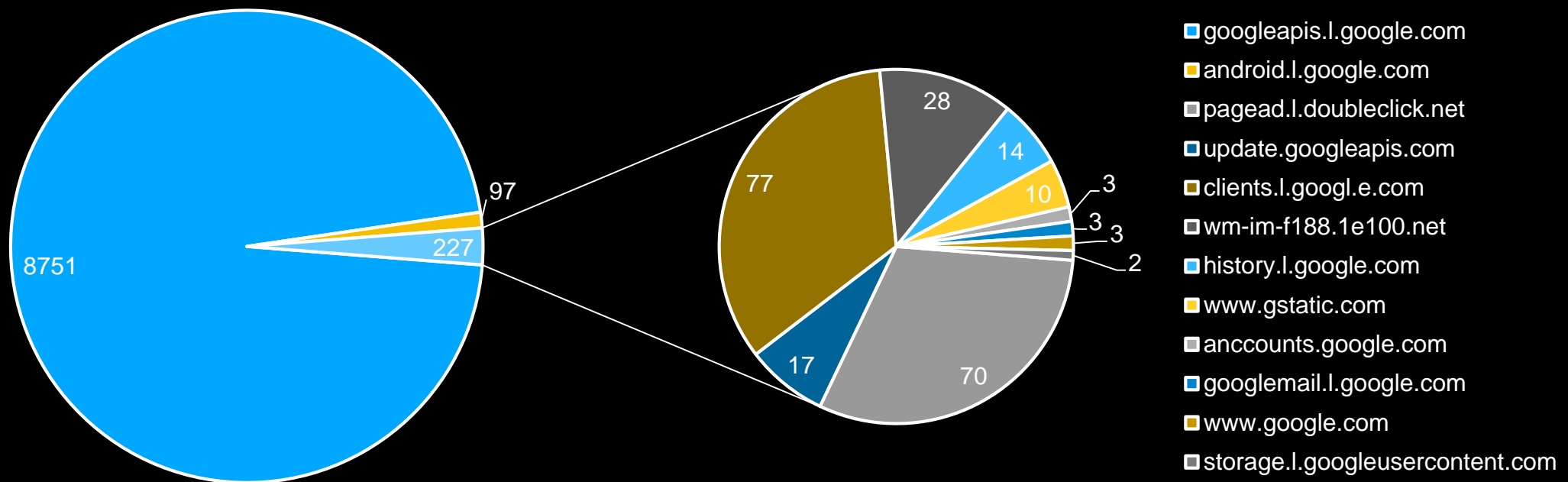
Packets Exchanged per Hour



Google Data Collection on Android

Some Data – My own analysis

Destination of Google Bound packets (per hour)



Google Data Collection on Android

Some Data – My own analysis

- HTTP Traffic: Location update:

Hypertext Transfer Protocol

```
POST /loc/m/api HTTP/1.1\r\n
User-Agent: GoogleMobile/1.0\r\n
Content-Type: application/binary\r\n
Transfer-Encoding: chunked\r\n
Host: www.google.com\r\n
Connection: Keep-Alive\r\n
Accept-Encoding: gzip\r\n
\r\n
[Full request URI: https://www.google.com/loc/m/api]
[HTTP request 1/1]
[Response in frame: 30554]
HTTP chunked response
File Data: 315 bytes
```

Media Type

Media type: application/binary (315 bytes) ← *The location object*

Google Data Collection on Android

Some Data – My own analysis

- HTTP Traffic: QuickSearchBox POST:

HTML Form URL Encoded: application/x-www-form-urlencoded

Form item: "androidId" = "3d4478c9dbb98ecb"

Form item: "lang" = "en_GB"

Form item: "google_play_services_version" = "11951446"

Form item: "sdk_version" = "25"

Form item: "device_country" = "gb"

Form item: "client_sig" = "38918a453d07199354f8b19af05ec6562ced5788"

Form item: "callerSig" = "38918a453d07199354f8b19af05ec6562ced5788"

Form item: "Email" = "darryncampbellfake@gmail.com"

Form item: "has_permission" = "1"

Form item: "service" = "oauth2:https://www.googleapis.com/auth/googlenow"

Form item: "app" = "com.google.android.googlequicksearchbox"

Form item: "check_email" = "1"

Form item: "token_request_options" = "CAA4AQ=="

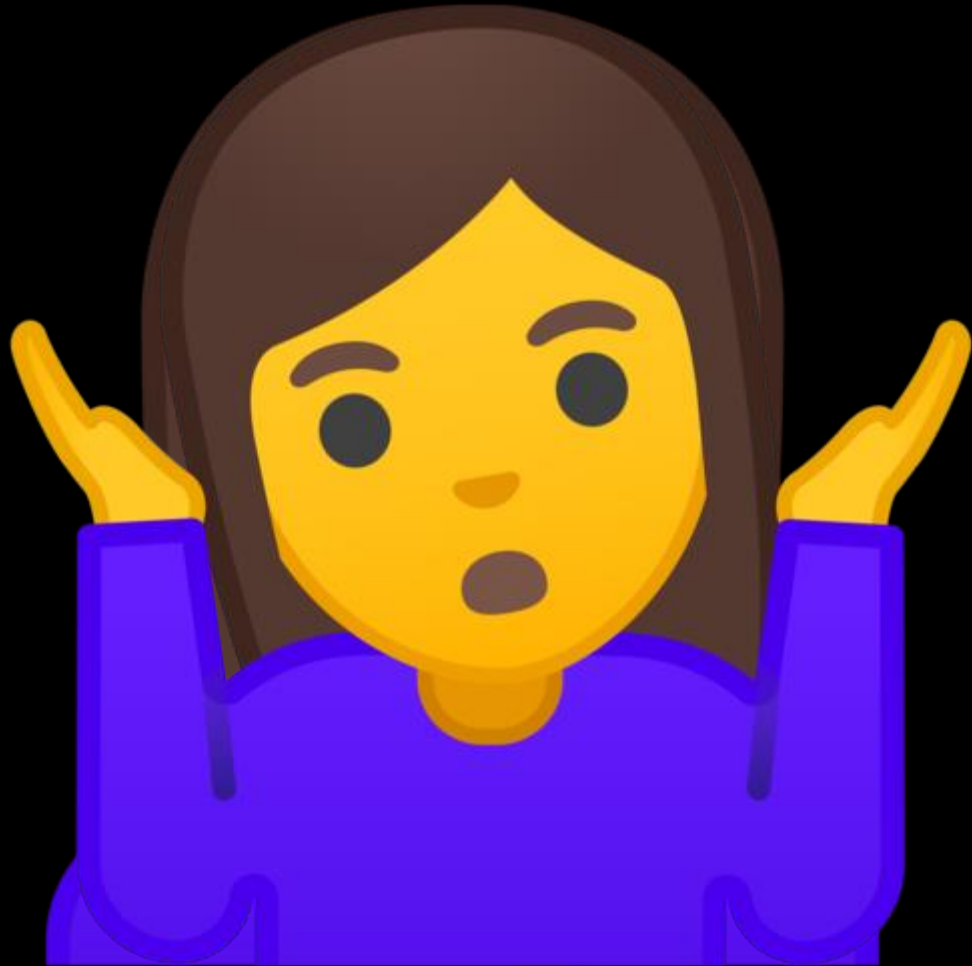
Form item: "system_partition" = "1"

Form item: "callerPkg" = "com.google.android.googlequicksearchbox"

Google Data Collection on Android

Some Data – My own analysis

- QUIC Data



Google Data Collection on Android

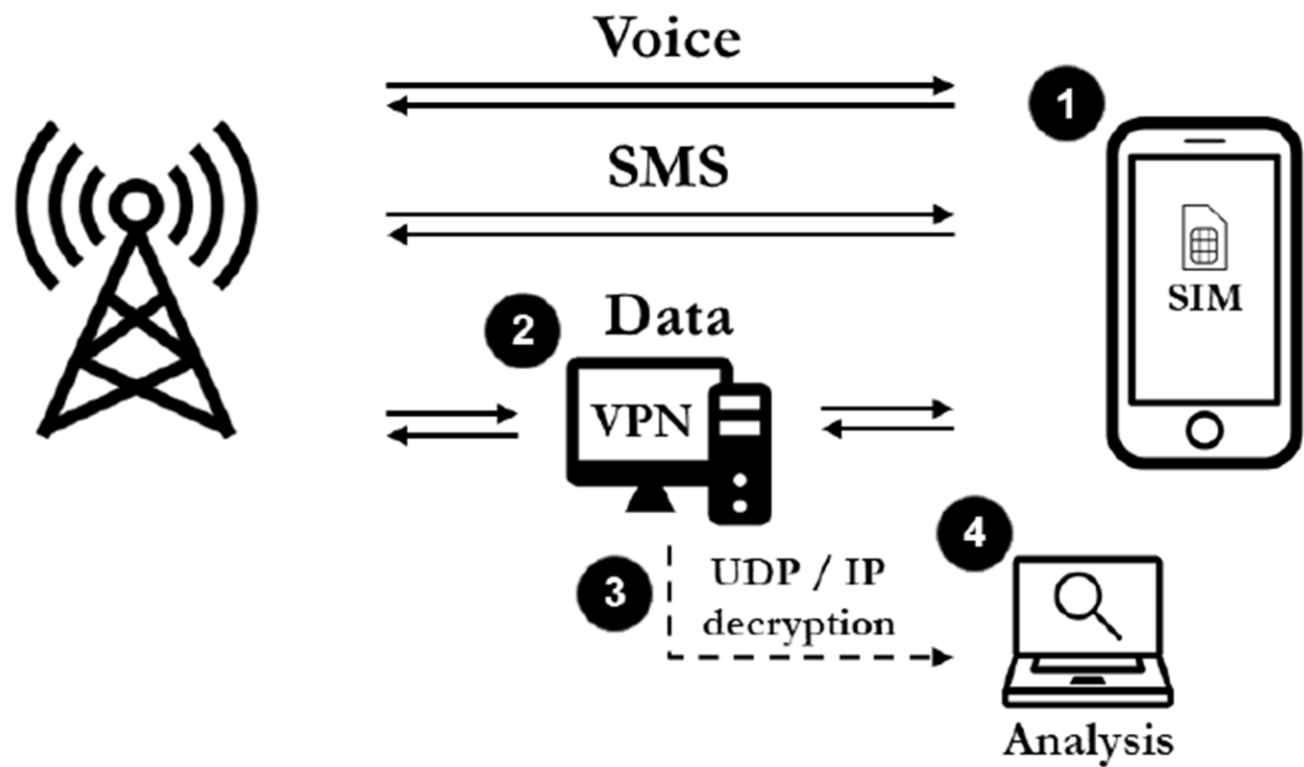
Enter Academia... Previous studies

Title	Relevant findings	Author, Date, Source
AP Exclusive: Google tracks your movements, like it or not	Google is tracking users' location even when location services are disabled	Ryan Nakashima August 2018 Source Link
Australian regulator investigates Google data harvesting from Android phones	Google "harvest" about 1GB of data from Android devices per month	Oracle May 2018 <i>Could not find original source</i>
Google collects Android users' locations even when location services are disabled	Google collects Android location when location services are turned off	Quartz November 2017 Source Link
Ericsson Mobility Report	Android devices consume more data (2.2GB/month) than other smartphones	Ericsson December 2013 Source Link
Data leakage from Android smartphones	Android passes anonymous IDs along with device IDs such as Mac addresses and IMEI	Lasse Øverlier June 2012 Source Link (Mirror)

Google Data Collection on Android

Performing your own tests

Figure 23: Example VPN setup used to analyze data shared with Google



ager (Oasis)

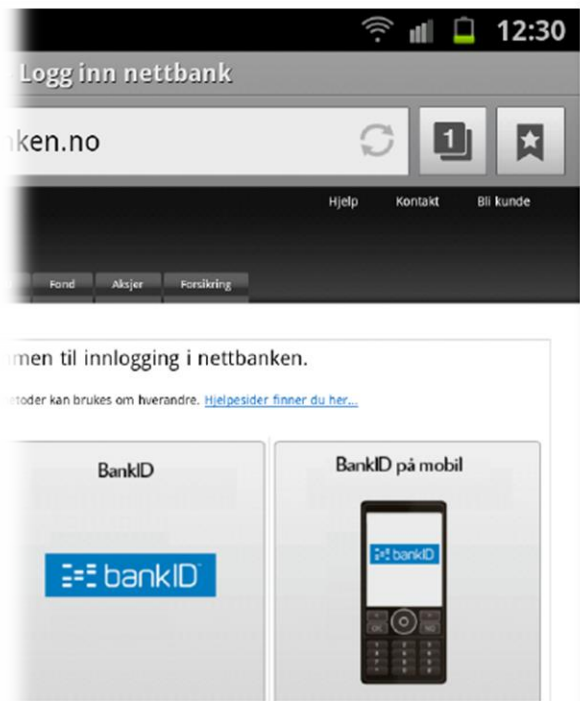


Figure 3.2 Login without warning when adding trusted certificate issuer locally

Google Data Collection on Android

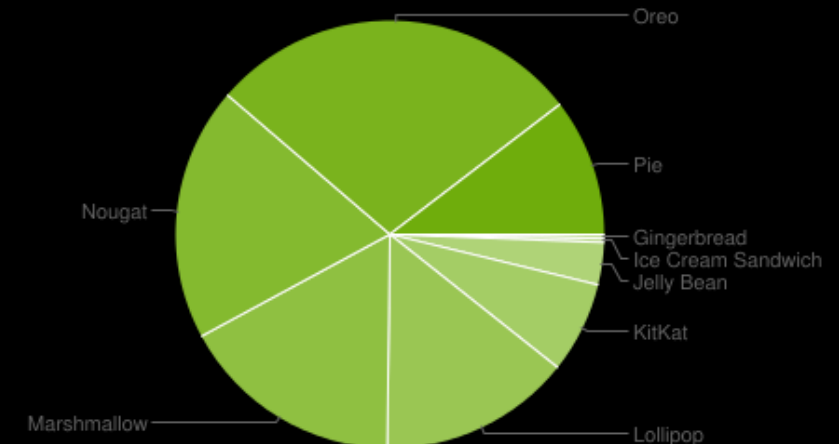
Enter Academia...

- August 2018, Professor Douglas C. Schmidt, Professor of Computer Science at Vanderbilt University, and his team published, “Google Data Collection”
 - “Offered to the public”
 - <https://digitalcontentnext.org/blog/2018/08/21/google-data-collection-research/>
- Collected data sources :
 - Google’s My Activity and Takeout tools
 - **Data intercepted as it is sent to Google server domains**
 - Google’s privacy policies
 - Other 3rd party research

Google Data Collection on Android

Some Data – Academic analysis

- Paper limits itself to Android 6.0 and (strangely) does not specify which phone is used. Android 6.0 has a 16.9% market share today.
- Paper conducts a series of tests on stationary and moving devices and determines typical usage based on past studies of user behaviour
- Paper distinguishes between ‘Active’ data collection and ‘Passive’ collection
- *“Overall, our findings indicate that Google has the ability to connect the anonymous data collected through passive means with the personal information of the user”*



Google Data Collection on Android

Some Data – Academic analysis

- *“Analysis... during a normal day in the life of Jane [test subject] suggested that the number of ‘passive’ data collection events outnumbered the ‘active’ events by approximately 2 to 1”*

Description	Active collection	Passive collection
Listening to Google Play in the morning	Music interests	Morning location
Walks from subway to work		Commute path to work address
Uses Maps to find a new lunch spot	Destination entered into Maps	Dining interests
Takes an Uber from home to work		Commute path home & use of the Uber app
Uses Google Home to play music for her children	Google Home search query	Location of Google Home

Google Data Collection on Android

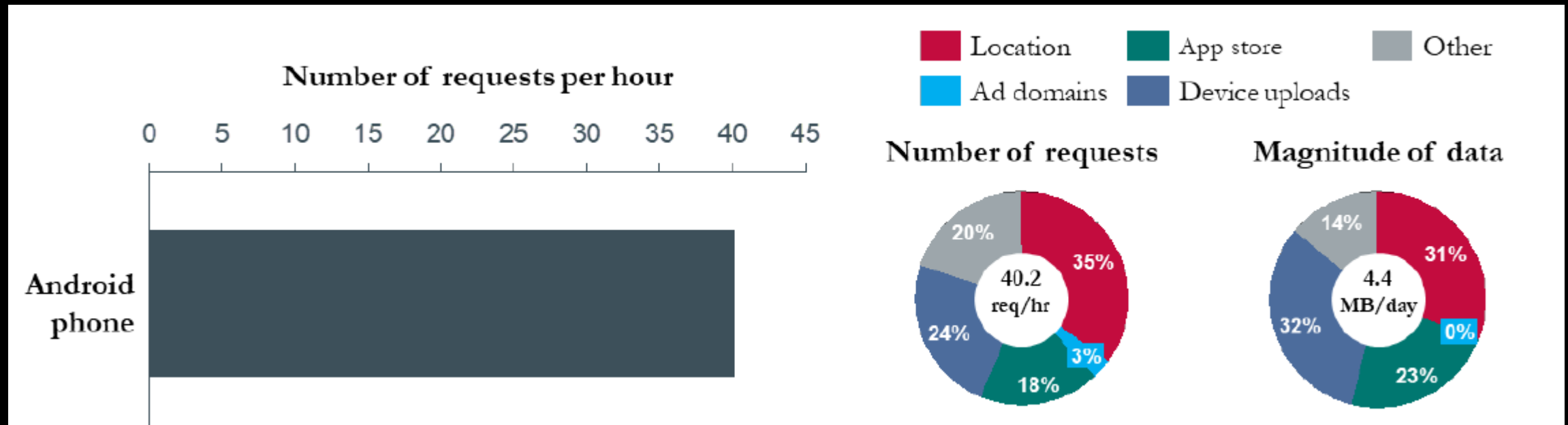
Some Data – Academic analysis

- Location
 - *“It is hard for an Android mobile user to “opt-out” of location tracking”.*
 - Location mode
 - Location Services
 - Wi-Fi
 - Wi-Fi scanning (even when Wi-Fi is off)
 - BLE scanning
 - Cell tower triangulation
 - Sensor dead reckoning
 - Google Maps location history
 - See: Recent ‘scare’ that Google is tracking your location when you turn location off (Quartz, Nov 2017)
 - *“During a short 15-minute walk around a residential neighbourhood, an Android device sent nine location requests to Google [containing] 100 unique BSSIDs of Wi-Fi access points”*

Google Data Collection on Android

Some Data – Academic analysis

- “Location constitutes most of the passive data sent from an idle device”



Google Data Collection on Android

Some Data – Academic analysis

- Location: Application APIs can return user activity using a combination of device sensors and on-device machine learning
- The academic analysis observed this data uploaded to Google

```
"activityReadings": [{  
  "activities": [{  
    "confidence": 99,  
    "type": "onFoot"  
  }, {  
    "confidence": 99,  
    "type": "walking"  
  }], "timestampMs": 1527095517507 },
```

Google Data Collection on Android

Some Data – Academic analysis

- Potential to de-anonymize device via data sent to Google by Android:
- The Check-in upload was observed 2 times in a 24 hour test:

Locale: "en-us"

macAddr: "2c598a23***"

MAC

Meid: "35956507040***"

IMEI

accountCookie: "[jane.gcollpro@gmail.com]"

GMAIL

accountCookie: "ya37.GokBuQWFvoYx919S7S52Hxksk-K0CZMUvVw1_hZO..."

...

serialNumber: "LGUS610b12c***"

Serial

...

Google Data Collection on Android

Some Data – Academic analysis

- Potential to de-anonymize device via data sent to Google by Android:
- The bulk upload allows the association between the GAID (Advertising ID) with the GSF ID (Unique to a user of Google services):

User-agent: "Android-Finsky/10.1.08-all%20..."

X-dfe-device-id: "3945dee97778***" GSF in hex

authorization: "GoogleLogin auth=lwZo4lcwvnmk8QKiuw..."

X-dfe-cookie: "CAEQARgAIjJVUygBMjxDaXNhS1FvVE5..."

X-def-device-checkin-consistency-token: "ABFEt1U16JQ5xikF9j8..."

...

X-ad-id: "83d479a5-5eb7-44f1-a2ee-14e91094" GAID

...

Google Data Collection on Android

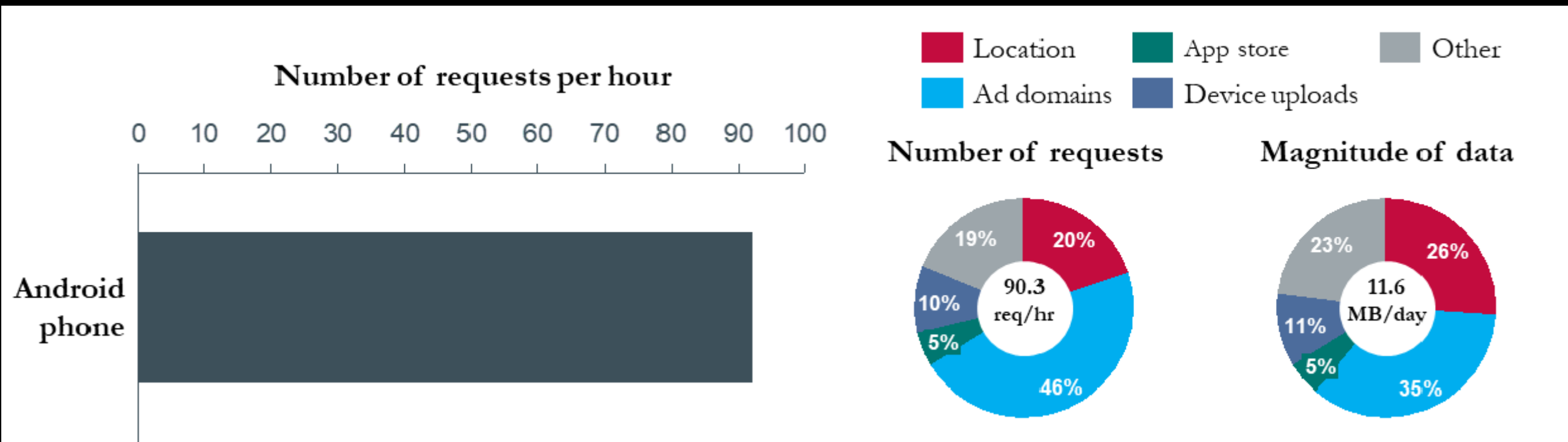
Some Data – Academic analysis

- Non-Stationary device
- “Various consumer usage studies were used to form a daily usage profile of a typical mobile phone user ... any direct interactions with Google’s products were omitted.”

Google Data Collection on Android

Some Data – Academic analysis

- Non-Stationary device
- Over a 24 hour period Google collected user location in ~450 instances, or 1.4x the experiment on a stationary device
 - This includes collection across all Google's portfolio of services e.g. DoubleClick, Analytics etc.

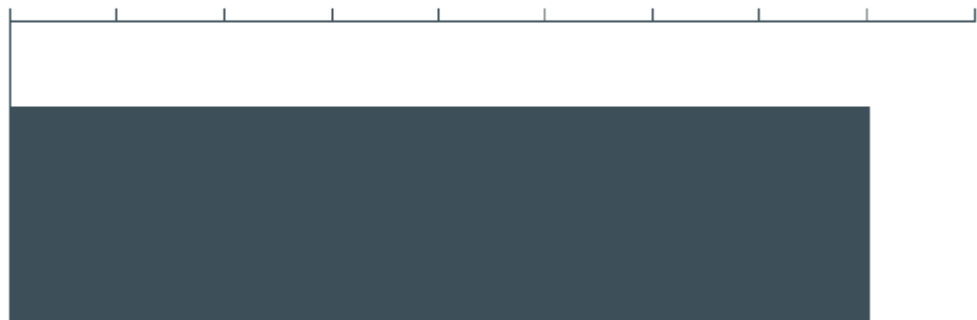




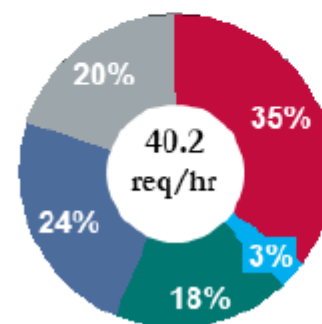
Number of requests per hour

0 5 10 15 20 25 30 35 40 45

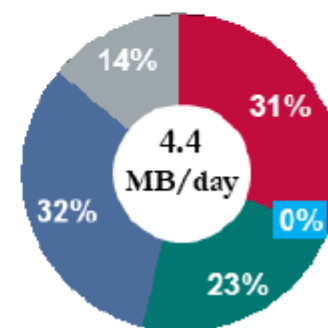
Android
phone



Number of requests



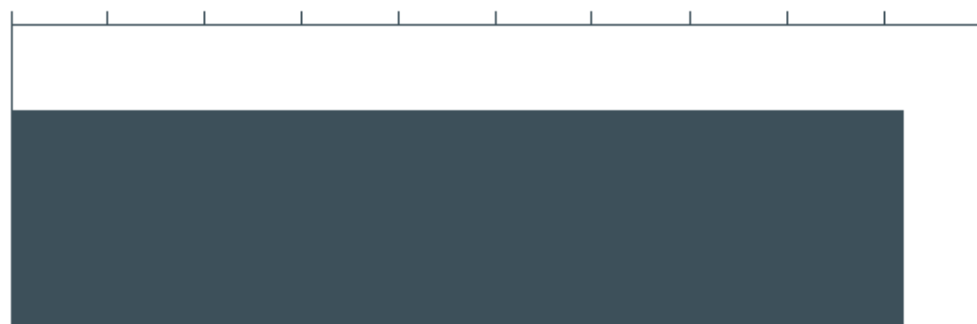
Magnitude of data



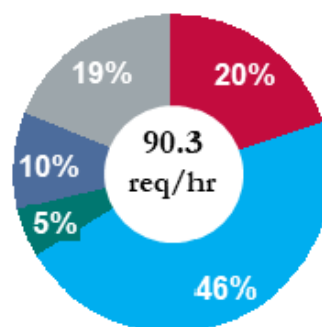
Number of requests per hour

0 10 20 30 40 50 60 70 80 90 100

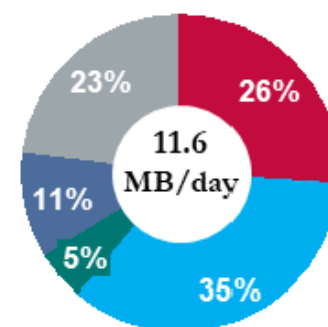
Android
phone



Number of requests



Magnitude of data



Google Data Collection on Android

Some Data – Academic analysis

- Conclusion:

- “A major part of Google’s data collection occurs while a user is not directly engaged with any of its products. The magnitude of such collection is significant, especially on Android module devices”
- “While such information is typically collected without identifying a unique user, Google distinctively possesses the ability to utilize data collected from other sources to de-anonymize such a collection”

Google Data Collection on Android

Some Data – Academic analysis

- Counterpoints:
 - None of these studies consider the value to the user of Google collecting this data
 - Consider the analogy with Runtime permissions: Why does Uber want to know my location? To provide me with a service.
 - There is an inherent assumption of nefarious purposes, less so in the academic studies but still present, *in my opinion*.
 - All of the studies analysed a very specific, narrow set of behaviour

Google Data Collection on Android

Some Data – Academic analysis

- As an aside...
 - “One specific request to Google that was further decoded was the ‘Google location API’ designated by the /loc/m/api endpoint. The location specifications were reverse engineered by removing the message header and decoding the compressed protobuf message”

Media Type

Media type: application/binary (315 bytes) ← *The location*

Google Data Collection on Android

Controlling your data

- Some dodgy websites out there giving all sorts of wacky advice on this!
- Some of the better ones I could find suggest:
 - Opt out of Ads personalization
 - Change DNS Server
 - Use a VPN
 - Don't use Google Apps
 - Disable app backup
 - Turn off Data & Personalization
 - Turn off Web & App activity
 - Turn off Location History

Google Data Collection on Android

Controlling your data

- Still a lot of closed source components, requires a degree of trust
- Every new version of Android introduces new components, settings, behaviour changes etc.
 - Trying to do this manually is a losing battle
 - You lose a lot of the value of Android by turning all this off
- Two approaches:
 - Network level
 - Device level

Google Data Collection on Android

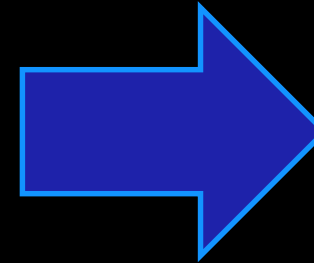
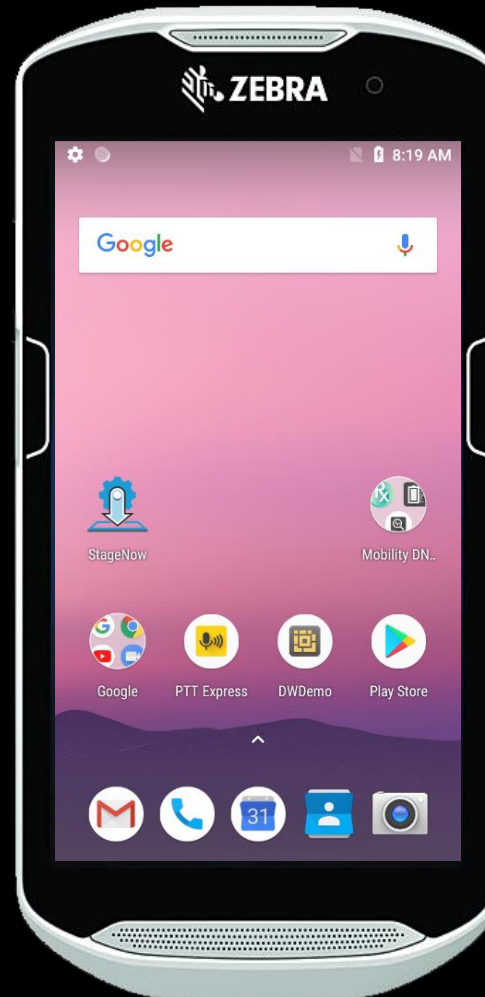
Controlling your data – Network Level

- No authoritative list of Google sites exists
- You COULD block all IP addresses listed under [AS15169](#)
 - Best source I could find for IP addresses owned by Google
 - Also includes 3rd party sites hosted by Google (about 8.7M in total)
- For 99.95% of people this approach would be a non-starter
- You COULD block named hosts based on previous traffic analysis
 - But you would need to re-analyse every time Play Store updates (not just Android)

Google Data Collection on Android

Controlling your data – Device Level

- Disable all GMS applications and services

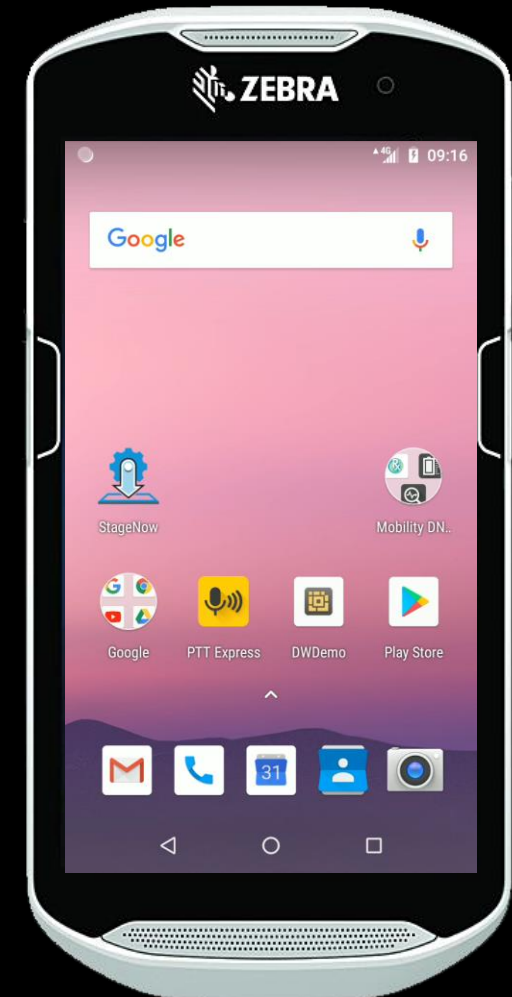


Google Data Collection on Android

Controlling your data

- Zebra's Restricted GMS provides an easy option for administrators to disable Google apps / services in a single step.
- Come talk to us at the booth!

- **DEMO**



Google Data Collection on Android

Resources

- News Organization articles:
 - [Guardian article](#)
 - [AP Exclusive: Google tracks your movements, like it or not](#)
 - [Google collects Android users' locations even when location services are disabled](#)
- My Original Research
 - [Network Stats Logger](#)
 - [Test Methodology](#)
- Academic Papers:
 - [Google Data Collection](#) (Aug 2018)
 - [Data leakage from Android smartphones](#) (June 2012)
- Controlling your Data
 - [Zebra's GMS Restricted](#) ([Webinar](#))

Google Data Collection on Android

Questions?

<http://developer.zebra.com>

Thank You



ZEBRA and the stylized Zebra head are trademarks of Zebra Technologies Corp., registered in many jurisdictions worldwide. All other trademarks are the property of their respective owners. ©2019 Zebra Technologies Corp. and/or its affiliates. All rights reserved.