SINTEF

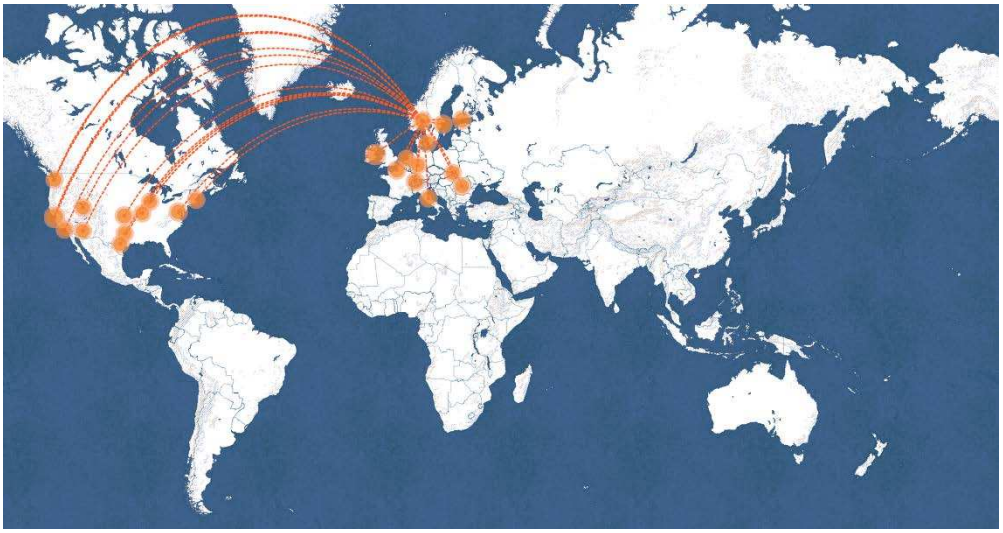
Report No. A27493

Report

Privacy in Mobile Apps

Measuring Privacy Risks in Mobile Apps

Author(s)
Antoine Pultier, Nicolas Harrand & Petter Bae Brandtzæg



SINTEF ICT
Networked Systems and Services
2016-01-06

SINTEF IKT
SINTEF ICT
Address:
Postboks 124 Blindern
NO-0314 Oslo
NORWAY
Telephone:+47
Telefax:+47 22067350

Enterprise /VAT No:
NO 948 007 029 MVA

Report

Privacy in Mobile Apps

Measuring Privacy Risks in Mobile Apps

KEYWORDS:
Privacy
Mobile Apps
Smartphones
Trackers

VERSION	DATE
Final	26.01.2016

AUTHOR(S)
Antoine Pultier, Nicolas Harrand & Petter Bae Brandtzaeg

CLIENT(S)	CLIENT'S REF.
Norwegian Consumer Council	Gro Mette Moen

PROJECT NO.	NUMBER OF PAGES/APPENDICES:
Project No. 102012440	19 + Appendices

ABSTRACT

Abstract heading

Privacy risks are increasingly linked to how people use their smartphones and tablets. This study investigates privacy issues in 21 mobile apps for Android. The experiment was done in Oslo, Norway, in November and December 2015. All the apps in this study accessed personally identifiable information.

A central finding is that many mobile apps not owned by big American tech companies (e.g. Google, Facebook) - such as sports apps and dating apps - transmitted potentially sensitive user data to a complex myriad of third-party services. In our study the 21 mobile apps communicated with approximately 600 different primary and third-party domains. Many of these third-party domains are trackers that pose potential privacy risks because we have little knowledge about how they collect, store and link user data. Third-party trackers in our study sent data to servers in Europe and the USA.

PREPARED BY	SIGNATURE
Petter Bae Brandtzaeg	

CHECKED BY	SIGNATURE
Jan Håvard Skjetne	

APPROVED BY	SIGNATURE
Bjørn Skjellaug	

REPORT NO.	ISBN	CLASSIFICATION	CLASSIFICATION THIS PAGE
A27493	9788214059250	Unrestricted	Unrestricted

Document history

VERSION	DATE	VERSION DESCRIPTION
V2.	2015 -12-23	Draft version submitted Norwegian Consumer Council
V3	2016-01-06	New version based on comments from Norwegian Consumer Council
Vfinal	2016-01-26	New version based on comments and proof-editing

Table of contents

1 Introduction

Privacy risks are increasingly linked to how people use the Internet and their smartphones. Internet privacy involves people's right to personal privacy concerning the storage, repurposing, provision to third parties, and display of information about oneself via the Internet. "Everyone has the right to respect for his private and family life, his home and his correspondence". (Article 8 of the Human Rights Act). Our right to privacy protects us against surveillance and intrusion into our personal lives.

A new challenge to people's privacy is that app services on our mobile devices, such as smartphones and tablets, are how these apps are collecting data and using intelligent tracking technologies to identify our interests, preferences, location, and health issues. According to Libert (2015, p. 1) "every new device, app, and social network is now assumed to come with hidden privacy risks". Automatic logging of user behaviour data reveals the applications and services we use, the articles we read, where we are in the world, what kind of music we listen to, the friends we have contact with and the interests we share with them, and so on. This intelligent information collection can filter content and tailor services to optimise the user experience, but at the same time, personal data is used by advertisers and others for unknown commercial purposes (Brandtzaeg & Lüders, 2009; Schmugar, 2008). Users are often unaware of such risks and their own privacy rights (Hoofnagle, Urban, & Li, 2012). However, there might be an increasing knowledge about these issues; a recent population survey in Norway shows that 76% have seen a connection between what they have searched online, and what advertising they receive. A total of 79 % believe it is uncomfortable that their personal data is collected, analysed and shared with other companies to show them customized advertising (Datatilsynet, 2016).

In digital societies such as Norway we see a usage pattern towards a mobile internet, where computer usage is to a small extent declining, and mobile smartphone usage is heavily increasing from 2012 to 2015 (see figure 1). According to Google's consumer barometer for 2015, 79% use a smartphone in Norway out of a representative population (online & offline) aged 16+.

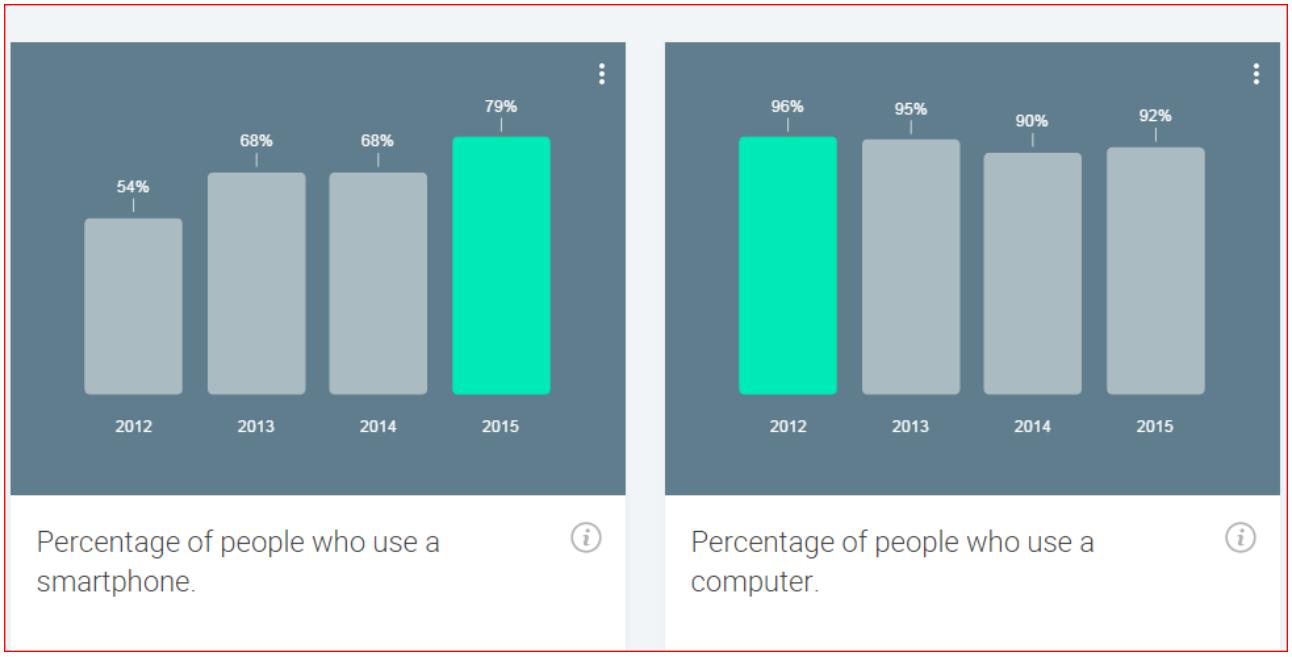


Figure 1. *Percentage of people using smartphones and computers in Norway from 2012-2015 (source: https://www.consumerbarometer.com)*

Norway is regarded as one of the most digital societies in the world, and the usage of apps on smartphones is high. The figure below shows what people mostly use their smartphone for in Norway:

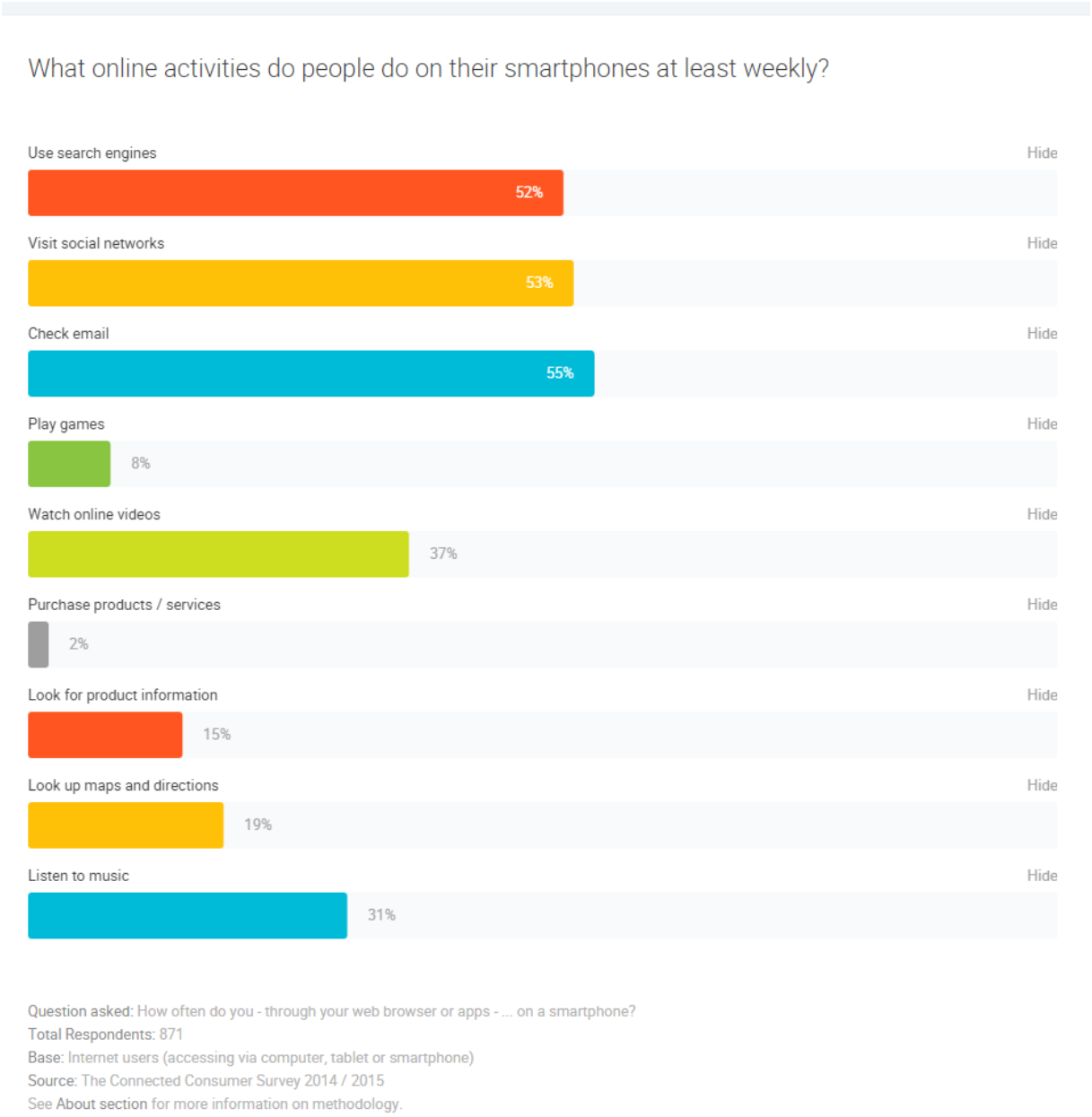


Figure 2. *What people use their smartphone for in Norway 2014/2015 (source: https://www.consumerbarometer.com)*

2 Aim

This study aims to provide an overview of what 21 different mobile apps that are popular among Norwegian smartphone users gather and reveal in terms of personal information. These apps cover some key application areas for mobile devices: online dating, social networking, gaming, news, entertainment, and money

transfers. The Norwegian Consumer Council selected the different mobile apps to be investigated based on their popularity among users in these categories. The apps are listed below[1]:

1. Facebook
2. Snapchat
3. Wordfeud Free
4. Tinder
5. Happn
6. Linkedin
7. VG
8. Endomondo
9. MyFitnessPal
10. Strava
11. Gulesider
12. Instagram
13. Lifesum
14. Messenger
15. Norli e-bok
16. Runkeeper
17. Twitter
18. Vipps
19. Yr
20. YouTube
21. Youtube Kids

3 Method

All the 21 mobile apps for Android in the list above were tested on two new smartphones, a Samsung Galaxy S5 and a Sony Xperia Z3 compact. The network access was a 4G connection provided by NetCom in Oslo.

This study collected and analysed data using a set of tools including Wireshark, Fiddler, Tcpdump, Ruby, SQLite3, TextQl, APKTool, Smali, Mechanize and PruneCluster. A laptop offered a Wifi access point. The laptop was running Wireshark and Fiddler to intercept the Samsung Galaxy phone's data. The Sony Xperia used tcpdump, requiring root access. These various tools were automated using a Ruby script. The same algorithm was used to automatically generate reports for each app (see appendix).

The HTTP and HTTPS communication between the app and application providers were particularly analysed, with the decoding of request parameters, cookies, HTTP headers, form submissions, JSON documents, and base64 encoded values. The Android application packages were decompressed and decompiled to detect the namespaces of the different software components, allowing us to list all the trackers. A more advanced analysis has also been used to detect what kinds of personal information are accessed by the different software components.

We tested the 21 mobile apps during the period from 12.11.2015 to 14.12.2015. All the mobile apps were tested from only one location: Oslo, Norway. One limitation is that the experimental setup and app tests were done over a short period of time, therefore, the results written in this report should be handled and interpreted based on this.

[1] We also included the FINN.no app in the study on a later stage (see results in appendix)

4 Results

4.1 Main findings

In this study we reveal many different third-party tracking services used by the majority of the mobile apps for Android that we have tested. Figures 3 and 4 illustrate a myriad of approximately 600 different primary or third-party domains, many of them third-party trackers, communicating with the 21 mobile apps we studied. Our results are similar to findings in a recent study published in Technology Science, where 73% of Android apps shared personal information, such as email addresses, with third parties (Zang, Dummit, Graves, Lisker, & Sweeney, 2015).

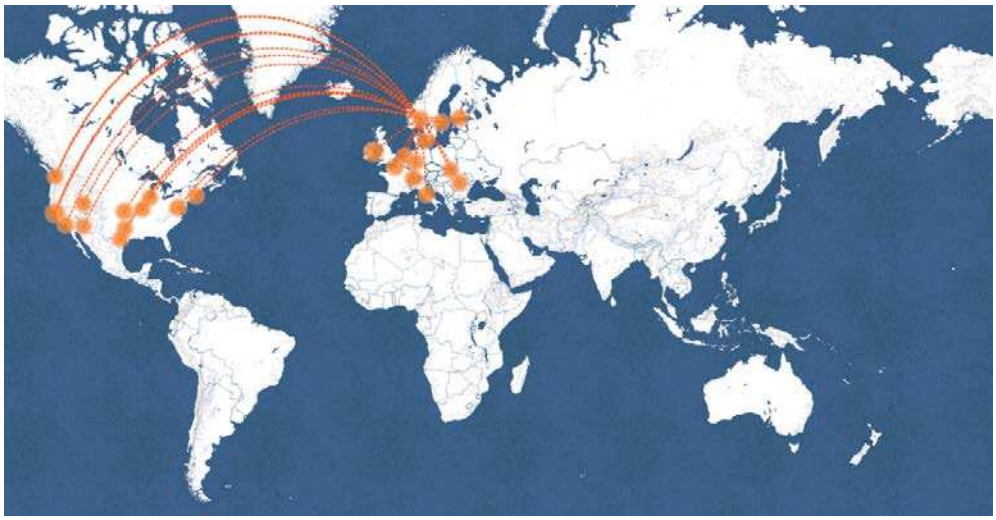


Figure 3. *The map illustration shows the geolocation of the contacted services for all the tested applications. The information from the 21 apps tested in Norway communicated with primary or third-party domains all over Europe and USA.*

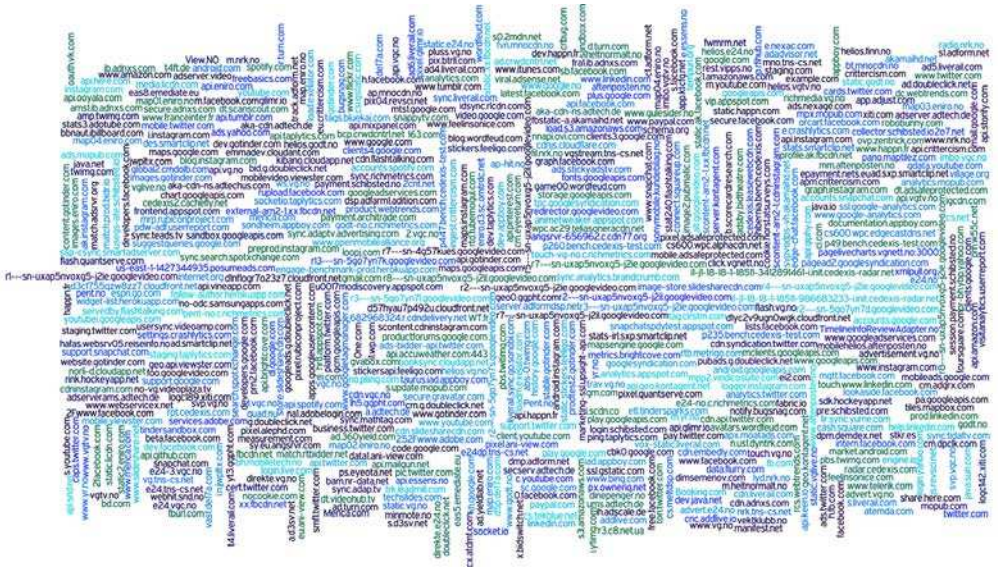


Figure 4. *This illustration shows the word cloud of approximately six hundred detected primary and third-party domains the 21 apps in this study communicated with. We expose a rich ecosystem of third-parties communicating with only 21 apps, which indicate the complexity in identifying where personal data are collected, stored and used.*

4.1.1 Third-party trackers as a privacy issue

A main finding is that mobile apps not owned by big American tech companies (e.g. Google, Facebook) - apps such as VG, sports apps and dating apps - often communicate and share personal information with many third-party trackers. This widespread sharing of user data with third-party services can potentially have large implications for people's privacy.

As illustrated in Figure 3, third-party trackers in our study sent data to servers in Europe and USA, but we know from other studies that third-party trackers use servers located all over the world (Falahrastegar et al., 2014). See also maps in the applications report for an overview of each app (see Appendix).

We can distinguish three categories of third-party trackers identified in this study:

1. **Advertising Trackers:** Targeted advertisers companies that sell ads based on user preferences and user location (e.g. Oslo). These trackers exist for the purpose of tracking and recording users' browsing habits, in addition to the tracking and recording of information on user behaviour patterns and interests, and to identify the user's particular device (e.g. the latest iPhone users spend more money than users of a cheap Android device).
2. **Application Monitoring Trackers**: These trackers monitor the usage patterns of the application. This gives the service provider data about the number of users in real-time, what the users are doing, where they navigate on the web, etc.
3. **Crash Reporting Trackers:** When an application crashes (a bug in the source code, an unexpected behaviour), a report is sent to help the developers analyse and fix the issue. The purpose of such trackers is to enhance the user experience and increase the security of the app.

If people are downloading and using mobile apps for Android such as Facebook, SnapChat, or the VG app, the users also reveal some personal data to the particular company hosting the app (e.g. Facebook or Amazon) as well as third-party trackers.

These kinds of tracked data are often about app users' activity (what kinds of sites they visit, for how long, what kind of content, such as video, images etc., they view) and personal information (gender, age, GPS location of their home/work place, device identifier to identify the same user across different applications, etc.). Trackers might sell these data to other companies and advertisers in one form or another.

In general, privacy issues and implications for users can be significant when all the data about user behaviour tracked by various third-party tracking services are linked or combined with personal information, such as email address, phone number, name IDs, IP address, cookie data, location, etc.

A major concern in regard to privacy risks is when users log into an app, if that app includes third-party advertising trackers, such tracking service can log how the users are using the app, identify which account in the app is owned by a particular user, etc. Consequently, third-party trackers can, if desired, go to a particular user app profile, record its contents, and add the contents to their own file of the user. Thinking of the younger user population, and the fact that 83% of children and young people between 9 and 16 years have smartphones (Barn & Medier, 2014), information about user profiles can be extensive when collecting and storing data about these young users over several years.

4.1.2 Applications and personal information

All the tested mobile apps in this study access personal information such as the device identifier, email, or other comparable data. Fourteen of the apps have permission to access the IMEI. We know that 4 of these 14 mobile apps send it in clear text or directly to their primary domains (Tinder, Vipps, Endomondo,

Runkeeper). We detected the personal information accessed by analyzing the code of the applications. Detecting access in the code does not prove that a particular app transmits personal information to others, but it is an indication that it might do so. By analyzing the data the network captured, we revealed that some mobile apps send personal information to third-party trackers. An in-depth code analysis revealed that personal information is often sent inside crash reports (although this did not occur during our tests).

Importantly most third-party trackers do not require personal information like name to identify users. IP addresses and a combination of device technical characteristics are often enough to compute device and user fingerprints. This data gathering likely allows tracking and advertising companies to identify the user across different accounts and applications (Boda, Földes, Gulyás, & Imre, 2010).

We know too little about what all the third-party trackers involved are doing with the collected user data. Small amounts of transmitted personal information is often not a problem related to user privacy, but the quantity and the frequency of such transmissions of user data, as we found in this study, make it problematic for user privacy when it is linked and combined. Most of the transmitted information is technical or generic, such as the device technical characteristics, the mobile phone operator name, the IP address, or the time zone. When summing and crosschecking this data it is possible to identify the user. Another problem is that users lack visibility and understanding of how trackers operate in regard to privacy. It is nearly impossible for users to see they are being tracked (Libert, 2015).

4.1.3 Cookies

A method for third-party trackers to submit data is to use cookies. Cookies and other methods which are similar to cookies are overabundant and are likely always used to identify the user, meaning the user is tracked, although we do not know to what extent. This was relevant for 17 of 21 apps (observed in the network captures).To reveal information about this we would need to set up a more comprehensive study. However, in contrast to web browsers, it is difficult or completely impossible for the mobile application user to manage her cookies. It also seems that third-party cookies are always accepted, allowing advertising services to identify and track the user. We detected that uninstalling the application Happn was not enough to remove its main cookie, which decreases privacy for the users of mobile apps. We can't tell if other apps do the same as Happn, so we need to test them for that.

On a general level, *crash report trackers* are documents containing as much information as possible to analyse when and why an application crashes. These trackers may therefore contain a lot of personal information that a user does not want to share. On the other hand, such information might be necessary for the developers to improve the user experience of the service.

4.1.4 American Tech Companies

Amazon, Google, Facebook, and content delivery networks, in general, have access to a large amount of information and could probably validate personal information between applications.
We find that these American tech companies have a major role in providing service, even for local Norwegian applications, such as YR. This pattern of data flow shows that data are only going between Western Europe and USA, not to other parts of the world such as China, Africa, or Russia. However, another study has shown that this was the case for many trackers (Falahrastegar et al., 2014).

4.1.5 Application security

Facebook, Messenger, Instagram and Snapchat detect the "man-in-the-middle" SSL data interception and refuse the connection. Therefore, we could not tell what these applications sent. From the user point of view, security is enhanced because the data cannot be stolen if the SSL chain of trust is broken.

4.2 Details for each app

4.2.1 Fitness/sport apps

In general, the tested fitness/sports applications contain many third-party trackers. This is, in particular, vulnerability in fitness and sports apps as they often contain health-related data. Collection of health data is a mega trend illustrated by the increasing use of smartwatches and bracelets that help users generate more statistics about their own body through sensors that measure heart rate, weight, movement, sleep, steps, etc[2]. Therefore the users of such apps will need clear information about how exactly the collected information will be used (e.g. Hamblen, 2015).

Moreover, the required Android permissions, which users accept, seem to bear little resemblance to the use or purposes of the many mobile apps. The majority of the fitness/sport apps sends the GPS-position of users not only back to the applications' own servers, but also to advertising companies. We did not detect data about body weight, size, or user performance being sent to third-party services.

Endomondo:

- Endomondo interacts with many third-party trackers. These third-party trackers have access to personal information, such as the IMEI number, location, and the device identifier.
- Endomondo sends the GPS position, age, and gender to Google's advertisement service and to a tracker (Rubicon Project).

MyFitnessPal:

- MyFitnessPal contains many third-party trackers and asks for unnecessary Android permissions (see applications report for details).
- These third-party-trackers have access to personal information about the phone and its user.
- MyFitnessPal sends the GPS-position to mopub.com for targeted advertising[3].
- MyFitnessPal synchronizes data even when the app and the phone are not in use (however, the phone needs to be on)(see the 48h analysis: see section 17).

2

In addition to share statistics of our health and movement patterns (e.g. dating and sports apps), a new coming trend might be emotional trackers, sharing the users' different emotions http://www.forbes.com/sites/jenniferhicks/2016/01/05/the-wearable-evolution-emotion-tracker-to-debut-at-ces-2016/?utm_campaign=ForbesTech&utm_source=TWITTER&utm_medium=social&utm_channel=Technology&linkId=20219677

3

As an example, we accessed their privacy policy 09.01.2016, the following information is written on their site http://www.mopub.com/legal/privacy/

LifeSum:

- LifeSum interacts with many third-party trackers. These trackers have access to important information about the phone and its user, such as the device identifier (Hardware Serial).
- LifeSum does not have access to the GPS position.

Runkeeper:

- Runkeeper contains many third-party trackers. These trackers have access to important information about the phone and its user, such as the device identifier.
- Runkeeper sends the GPS position to kiip.me, which is a third-party tracker.
- Runkeeper tracks the GPS position even when the phone is not in use (see the 48h analysis: see section 17).

Strava:

- Strava contains third-party trackers. These trackers have access to important information about the phone and its user, such as the IMEI and the device identifier.
- Strava sends the GPS position to its own servers.
- Strava publicly sells anonymised datasets.

4.2.2 Social networking sites apps

Facebook (1.59 billion users each month[4]):

- Most of the contacted Facebook servers are in Europe when accessing the app from Norway, Oslo.
- We were in this study unable to intercept the transmitted data. Facebook detects the MITM SSL data interception and refuses to connect. Except from a few Google domains (Facebook, Gmail, plus…) the majority of the detected domains are owned by Facebook.
- The source code is obfuscated, meaning we cannot easily determine what it does (not transparent). We have only detected the access of the device ID, but much more is accessed indeed.
- The Facebook app asks the user to accept many permissions, "some of these permissions sound scary" as Facebook (2016) themselves describe them. Most of them can, however, be explained by running the functionality and features of Facebook (see Facebook, 2016), and others are probably required due to technical reasons.
- Due to many permissions and the obfuscated source code there might be some privacy risks.
- The communications are encrypted. Facebook is very active also while the phone is not in use (see the 48h analysis), which indicates that the tracking continues also when the users are not using the

[4] Facebook Q4 and Full Year 2015 Earnings http://edge.media-server.com/m/p/r63ian8u

phone or the app. The application can be described as a "black box": you have no idea of what is inside, what it does, or when it does something.

- Critical permissions are: GPS position, read contacts, get accounts, record audio, read calendar, read SMS, read external storage, Wi-Fi access state (scan your Wi-Fi neighbourhood). However, Facebook can argue that those permissions are needed for running the app's functionality. Yet, those permissions are critical in regard to privacy and the user can't know how or when they are used.

Instagram (400 million users each month)[5]:

- Hosted by the European Facebook data centre, with content delivery network servers in northern Europe.
- No third-party trackers detected, only Facebook and Google are contacted. Instagram is owned by Facebook, and Facebook does not share information with third-party trackers to analyse their data directly. Hence, their business model relies on selling their own targeted advertisements.
- Instagram can access the users' phonebook and the important phone information such as GPS-locations. (See more details in the application reports.)

Messenger (800 million users each month)[6]:

The Messenger app is part of the Facebook service and is therefore also similar to Facebook, in regard to privacy issues. Hence, we were not able to reveal any transmitted data. The source code is obfuscated and the application asks for many Android permissions.

YouTube (over a billion users)[7]:

- YouTube communicates with Google's servers in California and the network provider Netcom's YouTube cache servers.
- A few cookies related to Google ads were detected.
- No third-party trackers were detected.
- In our experiment, we did not find that YouTube accesses much personal information, and the YouTube app will also ask the user for permissions.
- Similar to LinkedIn and Facebook, there are no third-party trackers. YouTube has its own data system. The YouTube data is of great value, and it makes sense for Google (YouTube) not to use another tracking system.

[5] Facebook Q4 and Full Year 2015 Earnings http://edge.media-server.com/m/p/r63ian8u
[6] Facebook Q4 and Full Year 2015 Earnings http://edge.media-server.com/m/p/r63ian8u
[7] YouTube statistics https://www.youtube.com/yt/press/statistics.html

YouTube Kids[8]:

- Similar to YouTube, but without Google ads, and fewer permissions from the user are required. However, Google does track kids, for example, through the operating system, Google Play, etc. Also, they use the channel for advertising towards kids, yet they have some restricted product categories (Youtube, 2016).

Twitter (320 million monthly users[9]*)*:

- The network communication is mainly sent to USA: Google's services and Google's cloud.
- Twitter can access the phonebook to find contacts, but asks the user for permission.
- Twitter can access the GPS by default, but asks the user for permission.
- Twitter has access to important personal information (device identifier, phone number), and can receive SMS and save data on the device. With the permission formulated and requested by Twitter, Twitter can read and write external storage, the application can also read/write data on the SD-card.

Snapchat (100 million daily active users)[10]:

- All the network communications go to Google's servers in California using an encrypted network tunnel.
- We were not able to see what went through this secured connection.
- The network security seems to be strong from the user's point of view; the data cannot be intercepted between Google's server and the device.
- A third-party tracker (flurry.com[11]) was revealed in the source code, but we cannot affirm whether it was used or not.

LinkedIn (400 million members)[12]:

- LinkedIn use a classic technological infrastructure with a LinkedIn data centre in USA and a content delivery network across Europe.
- No third-party tracker was detected, and only LinkedIn and Google were contacted.
- It can access the contacts, the phonebook, and information about the phone state and GPS.

[8] YouTube Kids is a new mobile app, launched in 2005.
[9] Twitter statistics: https://about.twitter.com/company
[10] Snapchat statistics http://expandedramblings.com/index.php/snapchat-statistics/
[11] Flurry embeds its software in several hundred thousands of apps on more than 1.2 billion devices to track usage. http://www.nytimes.com/2013/10/06/technology/selling-secrets-of-phone-users-to-advertisers.html
[12] About LinkedIn https://press.linkedin.com/about-linkedin

LinkedIn is a big social networking site and, similar to Facebook and YouTube (Google), they might not be dependent up on a third-party company to see what is going on in their system. But our results does not prove that LinkedIn resell user data.

4.2.3 Dating apps

Dating apps for mobile devices are quite new, but increasingly popular. Tinder (launched in 2012) and Happn (launched in 2014), investigated in this study, use GPS-location as an important feature that shows nearby users of the same app. According to Independent: "Happn works by tracking your location and noticing when you 'cross paths' with a fellow user, by coming within 250 metres of them. After crossing paths, you can then see their profile on Happn (and they can see yours), as well as the rough area where you passed each other and a few personal details" (Bolton, 2016). To use the Tinder and Happn apps users also will need a Facebook profile as both these apps connects with the user's personal Facebook profile. Hence, these apps encourage their users to share personal information and non-stop location data (GPS-location).

In this study we found a number of data types from these apps that raise concerns about user privacy. Similar to the sports/fitness apps, dating apps use mainly third-party trackers. Dating apps are also mapping users' GPS-location and personal identifiers. In more detail we find the following:

Happn:

- Happn contacts services located in Europe and USA, hosted by companies such as Online SAS, Amazon, Spotify, Google and EdgeCast.
- Happn stores a cookie that is not removed when uninstalling the application; consequently, a directory still remains on the user device containing some data. This implies that the users lose their ability to permanently remove the app and delete their personal information.
- Happn shares device identifier to a domain owned by UpSight, a major third-party tracking company, communicating very frequently about all user behaviour (e.g. liking other users, etc.) in the application. Every time users use the Happn app they are sharing information from their Facebook account, including name, age, birthday, job status, and gender, with a third-party tracker.
- Happn accessed and sent important personal information, such as the Facebook Identifier and the GPS position, very frequently. The use of GPS-location is well known and accepted by the user when they start using the service (The app considers that you have encountered someone if your paths cross between 1m and 250m)[13]. However, sending some of this information (age, name, gender, and Facebook identifier) to third-party tracking companies is probably less known. The GPS position is monitored while the phone is not used and is transmitted to Happn servers.

In addition to test the Android mobile app version of Happn, we also validated these results by using a same test procedure testing the iOS application in end of January 2016. We found similar results. The differences are the absence of servers hosted by Online AS in France, and that the technical details are specific to iOS instead of Android.

[13] Happn Faq https://www.happn.com/en/faq

Tinder (24 million members)[14]:

- Tinder interacts with several third-party trackers, and some of them access important personal information such as the device identifier.
- Tinder shares the GPS position with Taplytics, a monitoring/tracking service.
- Facebook and a few advertising services are also contacted.

It should be noted that a recent study found that the Tinder app is sharing more accurate location data than intended, as users could be located to within a small range of their present location (Dredge 2014).

4.2.4 Norwegian apps

Unlike the international apps, the Norwegian apps' servers are located in Norway and Sweden. We do not find the same excessive trend in terms of tracking and issues related to geolocation. VG is the worst in terms of the number of third-party trackers used. This may pose a problem to our ability to track and manage the different data that is now stored by different third-party trackers.

YR:

- Yr mobile app contains five third-party trackers. Crash analytics is the most active, probably necessary to optimize the user experience.
- No background activity happens and little network access is required. In general, using Yr is not found to be a threat to privacy issues, but some information is sent to third-party tracking companies.

Gulesider:

- This app has servers across the world – including in Sweden, Finland, Ireland, and the USA. The GPS coordinates are sent, but this make sense due to this being a map based application. SquareUp and LeadMill are two detected frameworks that could be considered as trackers.
- We detected problematic access to personal information, such as the Device ID and Sim card serial number. The contacts and call logs are also accessed, and it can receive SMS, but this may be explained by the purpose of the application.

WordFeudFree:

- This app contains third-party trackers that access personal information (such as the IMEI) and a few tracking services.
- We detected that WordFeud share GPS-location with a third-party tracker and advertising services.
- Some of the requested user permissions (e.g. GPS-location) are not required considering the purpose of the application, and are only present for the sake of tracking and targeted advertising.

[14] Tinder statistics 2015: http://www.datingsitesreviews.com/staticpages/index.php?page=Tinder-Statistics-Facts-History#ref-ODS-Tinder-2015-10

Vipps:

- This app contains two third-party trackers.
- This app can read the contact list and the system logs, which might be used for security reasons. However, system logs are critical information and it is not required for the general user functionality.
- Vipps contacts Facebook with a personal identifier when starting-up (we suspect the presence of a deactivated "connect with Facebook" button).

A relatively minor breach in the security was found testing Vipps and DNB (Den Norske Bank) was immediately notified. We considered the breach to not pose any large security threat to users of the app, but decided to notify DNB immediately, so that it could be fixed. DNB are currently working to fix this and the findings of our research will be published as soon as the the bug has been fixed, as to not put any users at additional risk. For further information please contact Kaj-Martin Georgsen at DNB.

VG:

- We detected several third-party trackers and advertisement services while using VG.
- The trackers are located mostly in Europe and USA, but spread across many countries and data centres.
- The GPS position is regularly sent, but it seems to be only used with the weather forecast service.
- We detected information being sent to several third-party-trackers. In more detail; the user behaviour (e.g. such as the users reading patterns and preferences) are shared with third-party trackers.

Norli ebook:

- This app uses Microsoft Azure cloud
- The synchronization services are called often (about 800 times in 48h). It is not necessary and uses too much battery.
- It contains one third-party tracker (Google Analytics).
- The app asked for permissions which are not required for the application's core purpose.

4.3 The 48 hours analysis

We installed all the applications on the Samsung Galaxy S5 smartphone and used them once. We connected our testing user accounts when required (login to Facebook, Snapchat, etc.). We then started the analysis and did not use the smartphone for 48 hours with the exception of checking the clock, but without unlocking the screen. The 48 hours analysis showed that a large majority of the network communications were with

Google's servers and Facebook's servers. Google provides notification services often used by third-party applications, which explains a part of the data flow. Norli e-bok uses a simple notification system, creating many network communications, but nothing important seems to be transmitted. As expected, the network activity was low and mostly used for notifications (85000 TCP packets totalling 19 Mbytes, an average 0.1kB/s).

We found that the GPS position is fetched and transmitted by Happn, which is one of the main application features. More surprisingly, we also detected that GPS location was sent by MyFitnessPal and Runkeeper when the applications are not in use. The users can then be geo-tracked whenever the GPS function is turned on.

5 References

Barn og Medier (2004). *En rapport fra Medietilsynet om barn og unges bruk og opplevelser av medier.* (A report on children and their use of media). Available online: http://www.medietilsynet.no/globalassets/publikasjoner/2015/rapport_barnogmedier_2014.pdf

Boda, K., Máté, A., Földes, G., Gulyás,G. & Imre, S. (2010). "User Tracking on the Web via Cross-Browser Fingerprinting" Available online: http://pet-portal.eu/files/articles/2011/fingerprinting/cross-browser_fingerprinting.pdf

Bolton, D. (2016). *Dating app Happn introduces new voice messaging feature.* Article in Independent, 20.01.2016. http://www.independent.co.uk/life-style/gadgets-and-tech/news/happn-voice-messages-online-dating-a6823936.html

Brandtzaeg, P. B., & Lüders, M. (2009). Privacy 2.0: Personal and consumer protection in new media reality. Oslo: Norwegian Consumer Commission. Available online: http://www.slideshare.net/PetterB/priv-20-rapport-om-sosiale-medier-og-personvern

Datatilsynet (2016). *Personvern – tilstand og trender 2016.* Available online: http://www.datatilsynet.no/Global/04_planer_rapporter/personvernrapporten-2016.pdf

Dredge, S. (2014). *Tinder dating app was sharing more of users' location data than they realised.* Guardian, 2014: Available online: http://www.theguardian.com/technology/2014/feb/20/tinder-app-dating-data-location-sharing

Facebook (2016). *Why is the Facebook app requesting permission to access features on my Android?* Available online: https://www.facebook.com/help/210676372433246

Falahrastegar et al., (2016). Tracking Personal Identifiers Across the Web. Available online: http://www.eecs.qmul.ac.uk/~hamed/papers/pam2k16.pdf

Falahrastegar, M., Haddadi, H., Uhlig, S., & Mortier, R. (2014). Anatomy of the third-party web tracking ecosystem. *arXiv preprint arXiv*:1409.1066

Googles Consumer Barometer (2015). Available online: https://www.consumerbarometer.com

Hamplen, M. (2015). Companies could use wearables to track employees' fitness, or even their whereabouts. *ComputerWorld.* Available online: http://www.computerworld.com/article/2925311/wearables/as-smartwatches-gain-traction-personal-data-privacy-worries-mount.html

Hoofnagle, C. J., Urban, J. M., & Li, S. (2012, October). *Privacy and modern advertising: Most us Internet users want'do not track'to stop collection of data about their online activities.* In Amsterdam Privacy Conference.

ITU (2015). *Global ICT data*: Available online: http://www.itu.int/net/pressoffice/press_releases/2015/57.aspx#.Vm81ArgrJD_

Libert, T. (2015). Exposing the Invisible Web: An Analysis of Third-Party HTTP Requests on 1 Million Websites. *International Journal of Communication*, 9, 18. Available online: http://ijoc.org/index.php/ijoc/article/view/3646

Schmugar, G, (2008)," The Future of Social Networking Sites" . Available online: http://www.wired.com/images_blogs/threatlevel/files/mcafee_security_journal_fall_2008.pdf

Youtube (2016). *Yotube help. Advertising on YouTube Kids:* Available online: https://support.google.com/youtube/answer/6168681?hl=en

Zang, K., Dummit, J., Graves, P.L, & Latanya S. (2015). Who Knows What About Me? A Survey of Behind the Scenes Personal Data Sharing to Third Parties by Mobile Apps. *Technology Science*. Available online: http://jots.pub/a/2015103001/

A Appendices

5.1.1.1.1.A.1 Results FINN

6

FINN.no:
FINN.no AS is a Norwegian online marketplace, which conveys different types of products and services. FINN enables purchases and sales between private individuals and between private individuals and companies. FINN.no is one of Norway's most popular network services measured by the number of unique users. Measured by the number of page views is FINN.no the largest Norwegian service on-line. Overall, it was published in excess of 5 million ads on FINN in 2014[15].

To test FINN, we used the same method as described in the report, but in late January and not in December.

Results:
We detected six different tracking and targeted advertisement components. The third-party tracker "omtrdc.net" from Adobe in USA, and the primary tracker "collector.schibsted.io"[16] from Schibsted and hosted by Amazon in Ireland are the two most active. The targeted advertisement services were Adtech.de from AOL and accessed through Akamai CDN network (the servers were located in Switzerland and USA in these tests), Google Ad Services and Facebook. In these tests, we did not use Facebook related features. However, we classify it as a tracker and targeted advertisement because we captured a request initializing the Facebook component with tracking options (adverstiser_tracking_enabled and application_tracking_enabled).

By analysing the application code, we can see the device identifier accessed by spring.de (the tracker related to Kantar), but we did not detect it being sent. The network operator name is accessed by Adobe's tracker and Facebook. Adobe's tracker (previously named Omniture) and Schibsted's tracker log the application usage, e.g. the consulted ad. The test procedure included the FINN feature to find ads in the user area. It appears the GPS position is communicated with FINN's servers, and the Schibsted tracker.

FINN uses few strong identifiers about the phone hardware and the application version. For example, the application explicitly declares its web user-agent with "UA spoofed for tracking". These identifiers are probably required to help the development process to identify technical problems. FINN also uses a cookie named "unique visitor id".

Main FINN's servers are located in Oslo and overall the application communicates with servers located in USA, and north of Europe.

6.1.1.1.1.A.1 Html reports

All the data can be further analysed in the reports archive. Each application has a dedicated folder containing the report named index.html. The following is a short summary of the detected transmitted data.

IMEI, SIM ID: Tinder, Happn, Vipps, Endomondo, Gulesider

[15] Finn.no: https://no.wikipedia.org/wiki/FINN.no
[16] Schibsted is the company owning Finn.no. We classified the service "collector.schibsted.io" as a tracker considering its namespace in the application code (com.shipsted.spt.tracking) and the transmitted data. It is, however, not a third-party tracker.

GPS position: VG, Endomondo, MyFitnessPal, Strava, Yr, Happn, Gulesider, Tinder, WordFeudFree, RunKeeper

Network mobile provider: Tinder, Happn, WordfeudFree

For convenience, the HTML reports are also hosted temporary on a SINTEF server. This is the list of the applications reports:

http://forbrukerraadet.master-bridge.eu/output/endomondo/
http://forbrukerraadet.master-bridge.eu/output/facebook/
http://forbrukerraadet.master-bridge.eu/output/gulesider/
http://forbrukerraadet.master-bridge.eu/output/happn/
http://forbrukerraadet.master-bridge.eu/output/instagram/
http://forbrukerraadet.master-bridge.eu/output/lifesum/
http://forbrukerraadet.master-bridge.eu/output/linkedin/
http://forbrukerraadet.master-bridge.eu/output/messenger/
http://forbrukerraadet.master-bridge.eu/output/myfitnesspal/
http://forbrukerraadet.master-bridge.eu/output/norliebook/
http://forbrukerraadet.master-bridge.eu/output/runkeeper/
http://forbrukerraadet.master-bridge.eu/output/snapchat/
http://forbrukerraadet.master-bridge.eu/output/strava/
http://forbrukerraadet.master-bridge.eu/output/tinder/
http://forbrukerraadet.master-bridge.eu/output/twitter/
http://forbrukerraadet.master-bridge.eu/output/vg/
http://forbrukerraadet.master-bridge.eu/output/vipps/
http://forbrukerraadet.master-bridge.eu/output/wordfeudfree/
http://forbrukerraadet.master-bridge.eu/output/youtube2/
http://forbrukerraadet.master-bridge.eu/output/youtubekids/
http://forbrukerraadet.master-bridge.eu/output/yr/
http://forbrukerraadet.master-bridge.eu/output/finn/

48h report:

http://forbrukerraadet.master-bridge.eu/output/toutappv2/

The file personal.html contains all the detected data in each experiment.

http://forbrukerraadet.master-bridge.eu/output/personal.html

6.1.1.1.1.A.2 Notes on the reports

Dangerous permissions:

We used the Google Android documentation to determine whether a system permission is dangerous.
http://developer.android.com/guide/topics/security/permissions.html#normal-dangerous

Personal Information:

We developed a list of permission susceptible to access personal information. The full list is the following:
`ACCESS_COARSE_LOCATION, ACCESS_FINE_LOCATION, ACCESS_LOCATION_EXTRA_COMMANDS, ACCESS_WIFI_STATE, ACCOUNT_MANAGER, BIND_DEVICE_ADMIN, BIND_NFC_SERVICE,`

BIND_VPN_SERVICE, BLUETOOTH, BLUETOOTH_ADMIN, BLUETOOTH_PRIVILEGED, BODY_SENSORS, CALL_PHONE, CALL_PRIVILEGED, DUMP, GET_ACCOUNTS, GET_ACCOUNTS_PRIVILEGED, GET_TASKS, GLOBAL_SEARCH, MANAGE_DOCUMENTS, READ_CALENDAR, READ_CALL_LOGS, READ_CONTACTS, READ_EXTERNAL_STORAGE, READ_LOGS, READ_PHONE_STATE, READ_SMS, READ_VOICEMAIL, RECEIVE_MMS, RECEIVE_SMS, RECORD_AUDIO.

Detected programmatic access

The detected programmatic access consisted of automated analysis of the application bytecode. It provides an overview of what an application contains and access. However, the detected bytecode could be dead-code, or the application code could be obfuscated and important programmatic access may remain undetected.
We classified the programmatic access in the following categories:
Camera, Contacts, Location, Microphone, Call logs, Read SMS, Body Sensors, IMEI and phone states. Because the phone states consists of various access to technical or personal information, we used three subcategories. Light for insignificant technical information, Average for technical information that is not personal, and Strong for personal information.

PhoneStatesAverage: getNetworkCountryIso, getNetworkOperator, getNetworkOperatorName, getSimCountryIso, getSimOperatorName

PhoneStatesStrong: getCellLocation, getAllCellInfo, getDeviceId, getGroupIdLevel1, getLine1Number, getNeighboringCellInfo, getSimSerialNumber, getSubscriberId, getVoiceMailAlphaTag, setPreferredNetworkTypeToGlobal, onCellLocationChanged, onCellInfoChanged

Detected information

The detected information is a summary of all the information we automatically analysed in the recorded data. Technically, we parsed or decoded HTTP GET parameters, HTTP cookies, HTTP headers, HTTP POST requests, JSON documents and Base64 encoded values. For each value in the table, we list each related name, host, source and type of detection. The table is sorted by frequency. We manually searched important values in the table, using the web browser search feature. The values to search are context dependant and require some exploration. However, we used the following list as a starting point. We also used context dependant keywords such as "weight", "gender", and "age" when needed.

Name to search	*Value to search*	*Description*
Lon	10.7	GPS position
Lat	59.9	GPS position
android_id	4DE05342FA047907	Android Device ID
Device-id or android-id	3688FA76D9AA6BF1	Google Service Framework
IMEI	353008074731783	IMEI number
Serial	33009139c24b62b1	Hardware Serial
Sn	RF8G91FTKQY	Serial number
Fbid	111034575929723	Facebook ID
.id or _id	*	Generic identifiers

Table 1: important values to search

The map

The map in each of the html reports shows the location of the contacted servers while the applications were tested. We computed this data using the TCP/IP captures. The servers are located from their IP address using GeoLite2 from MaxMind (released under the Creative Commans Attribution-ShareAlike 3.0 license).

The orange circles are clusters of TCP/IP packets. The cluster numbers represent the number of TCP/IP packets. You can zoom in and zoom out on the map to change the cluster area, or to have a more precise view or an overview. A click on the cluster will show the detected hosting companies for the cluster area.
The test duration is displayed on the map bottom left.

Technology for a better society

www.sintef.no