Privacy issues in Android applications: the cases of GPS navigators and fitness trackers

Stylianos Monogios, Kyriakos Magos and Konstantinos Limniotis*

School of Pure and Applied Sciences,
Open University of Cyprus,
G. Kranidioti Ave., 2220, Nicosia, Cyprus
Email: monostelios@gmail.com
Email: kyriakosmagos@gmail.com
Email: konstantinos.limniotis@ouc.ac.cy
*Corresponding author

Nicholas Kolokotronis

Department of Informatics and Telecommunications,
University of Peloponnese,
Akadimaikou G.K. Vlachou Street, 22131 Tripolis, Greece
Email: nkolok@uop.gr

Stavros Shiaeles

Cyber Security Research Group,
School of Computing
University of Portsmouth, Portsmouth PO1 2UP, UK
Email: sshiaeles@ieee.org

Abstract: The Android operating system constitutes a very attracting platform for developing smart applications providing various services to the users, including the field of e-governance. This provision comes along with personal data processing, which in turn raises several privacy concerns. This paper studies privacy issues in the mobile ecosystem, focusing on two important types of smart applications which process personal data to a large extent: global positioning system (GPS) navigators and fitness tracking applications. More precisely, for both types of applications, an indicative list of popular apps is being analysed through appropriate experimental environment, aiming to identify the underlying personal data processing that takes place. Our analysis illustrates that both GPS navigation apps and fitness trackers have access to several types of users data, while they may allow for personal data leakage towards third parties such as library providers or tracking services without providing always adequate or precise information to the users.

Keywords: Android system; fitness trackers; GPS navigator; privacy; profiling; third-party library.

Reference to this paper should be made as follows: Monogios, S., Magos, K., Limniotis, K., Kolokotronis, N. and Shiaeles, S. (2022) 'Privacy issues in Android applications: the cases of GPS navigators and fitness trackers', *Int. J. Electronic Governance*, Vol. 14, Nos. 1/2, pp.83–111.

Biographical notes: Stylianos Monogios received his BSc as a Deck Officer from a Supreme Military Educational Institution, the Hellenic Naval Academy (HNA) in 2013, and MSc in Information Systems Security from the Open University of Cyprus in 2019. He is currently working as a Sub-Lieutenant in the Naval Command of Cyprus National Guard. His research interests include personal data protection and information security.

Kyriakos Magos received his Master's degree in Information Systems Security from the Open University of Cyprus in 2020. He is currently working in Cyprus National Guard as a Lieutenant. His research interests include cryptography and personal data protection.

Konstantinos Limniotis received his PhD in Cryptography from the National and Kapodistrian University of Athens in 2007. He is currently an Adjunct Faculty member at the Open University of Cyprus, as well as an ICT Auditor at the Hellenic Data Protection Authority and a Research Associate with the University of Peloponnese, Greece. His research interests include cryptography, personal data protection, information security and coding theory. He has more that 40 publications (book chapters, refereed international journals and conferences) in the above areas.

Nicholas Kolokotronis is an Associate Professor and head of the Cryptography and Security Group at the Department of Informatics and Telecommunications, University of the Peloponnese. He received his BSc in mathematics from the Aristotle University of Thessaloniki, Greece (1995), MSc in highly efficient algorithms (1998, highest honors) and PhD in cryptography (2003) from the University of Athens. He has been a cochair and member of the organising committee in many international conferences. His research interests span cryptography, security, and coding theory, areas in which he has published more than 85 papers in international scientific journals, conferences, and books.

Stavros Shiaeles is a currently Senior Lecturer at University of Portsmouth, UK. He holds various professional certifications in Cyber Security and he is an EC-Council Accredited Instructor, delivering CEH to professionals. He has published more than 70 papers in international scientific journals, conferences and books, he has been a TPC member and a regular reviewer for a number of international journals and conferences, and he is actively involved with EU-funded and national research and development projects. His research interests span the broad areas of cybersecurity, open-source intelligence, trust, blockchain, digital forensics, and machine learning applications in cybersecurity.

This paper is a revised and expanded version of a paper entitled 'A case study of intra-library privacy issues on Android GPS navigation apps' presented at the *8th International Conference on e-Democracy (www.edemocracy2019.eu)*, Athens, 12–13 December 2019.

1 Introduction

The smart mobile applications ecosystem has met tremendous developments during the last years, being still though a highly evolving environment. According to publically available statistical information (Statista, 2020), the number of smartphone users in the world in 2020 was 3.5 billion or, equivalently, about 44.98% of the world's population owns a smartphone. This constitutes a rapid growth during the last years, since in 2016 there was 2.5 billion users – i.e., an increase by 33.58% has been occurred during the last 4 years. Regarding the underlying platform, smartphones running the Android operating system held an 87% share of the global market in 2019 (Statista, 2019). The types of smart applications span several diverge fields, such location-based services, entertainment, e-commerce services, e-banking services etc., whereas also applications in the area of health are being used, being referred as mHealth (which is a general term for describing the use of smart mobile phones in medical care). Smart applications are also being widely used in the context of e-governance (see, e.g., Pang (2018)). Bearing also in mind that Internet-of-Things (IoT) solutions (platforms and services) can also be accessed via mobile apps, as well as that the next generation of mobile networks technology will realise part of the IoT's connectivity, it becomes evident that smart applications are still expanding.

In this complex environment of the smart mobile applications, personal data protection constitutes an important challenge from both technical and legal aspects. Indeed, smart applications may process large amounts of personal data, such as the users' location, friendships, habits, interests or even health data – thus developing profiles of the users. This information can be used for commercial purposes, including behavioural advertising, although it may go far beyond this purpose – e.g. to infer a user's socio-economic class, health status or political beliefs. Such privacy issues are further accentuated by the fact that machine learning – a form of artificial intelligence – is also rapidly growing, allowing machines use data to learn on their own, with the ultimate goal to strengthen their capability of deriving safe conclusions based on (big) data analysis – e.g., from predicting what customers want to buy to identifying people at risk for a certain disease or their personality in the context of political campaigns (Chester and Montgomery, 2017; Mavriki et al., 2019).

Generally, several tracking mechanisms of different forms exist (Castellucia, 2012; Bujlow et al., 2017), aiming to create profile of the users. For example, towards implementing behavioural advertising, (efficient) tracking mechanisms is a prerequisite for the ad networks. Probably the most difficult one to be tackled towards protecting users' privacy rests with the generation of a so-called fingerprint of the user – that is, a unique identifier of a device, operating system, browser version, or other instance that can be read by a web service when the user browses, allowing the tracking of the user when he visits several websites belonging to different entities. Fingerprinting was first defined as *browser* fingerprinting in Eckersley (2010) and has been subsequently generalised to describe any unique instance that a device leaves, which can be based on, e.g., a specific software that is installed on the device or the particular device configurations (Kurtz et al., 2016). The difficulty in dealing with fingerprinting rests with the fact that fingerprints are not based on any client-based storage (such as the case of cookies) and thus sophisticated *data protection by design* solutions are needed to alleviate the relevant privacy risks. Especially in the mobile applications ecosystem, behavioural advertising can be upgraded into ubiquitous advertising (Krumm, 2010), that is advertisements will not only be personalised to users' online profiles, but also to their physical profiles – e.g., advertisements will be customised to users' locations, physical or intellectual activities, etc. (see Castellucia (2012)).

The average smartphone has more than 25 apps installed (see Taylor et al. (2017)), each having its own access rights to the device depending on the permissions that the user grants. The vast majority of the apps utilise third-party libraries for several purposes – e.g., to provide integration with social networks or to facilitate the programming procedure via easily embedding complex functionalities. These libraries obtain the same access rights with the host app. However, the use of such libraries may pose some risks for the users' privacy, since they may, e.g., track the users (Stevens et al., 2012; Binns et al., 2018). Moreover, as it is analysed in Taylor et al. (2017), the use of several popular libraries by several different smart apps may result in the so-called *intra-library collusion*, that is the case that a single library embedded in several apps on a device may appropriately combine the set of permissions given by each host app so as to leverage the acquired privileges and gather (a possibly large amount of) personal information without the explicit consent of the user. More specifically, as also stated in Taylor et al. (2017), the current Android security model, which does not support the separation of privileges between apps and the embedded libraries, facilitates the following relative privacy threats without the user's consent:

- libraries may abuse the privileges granted to the host applications
- libraries may track the users
- libraries may aggregate multiple signals for detailed user profiling.

More than half of the apps available on Google Play contain ad libraries linked to third party advertisers (Athanasopoulos et al., 2016). According to Ren et al. (2018), where several versions of popular Android apps have been examined in terms of whether privacy issues are being efficiently addressed over time, there is still an increased collection of personally identifiable information across app versions, a slow adoption of HTTPS to secure the information sent to other parties, and a large number of third parties being able to link user activity and locations across apps. According to several studies, users choose to install ad blockers in order to improve their user experience, to achieve better performance as well for security and privacy protection purposes (see e.g.,Mattke et al. (2017)). The work of Gervais et al. (2017) illustrates that, although the use of ad blockers provides a significant improvement in terms of user privacy, the degree of provided protection is highly depended on their configuration. A comparative study between several web privacy protecting techniques is given in Mazel et al. (2017). Interestingly enough though, in Icram and Kaafar (2017) it is shown that even in privacy enhancing technologies such as ad blockers we may encounter privacy issues, since the analysis therein indicates that neither ad blockers are free of third-party tracking libraries and permissions to access critical resources on users' mobile devices.

This paper focuses on the privacy issues in the Android ecosystem, putting emphasis on two specific types of smart applications: global positioning system (GPS) navigators and fitness tracking applications. These types of applications are of special nature, since the first one necessitates access to the current device's geolocation data, whereas the second processes data that could yield sensitive personal information related to user's health. Our approach is based on analysing, for both cases, the user's personal data that such applications process and examining whether this process may pose some (hidden) risks for user's privacy. In this direction, we studied popular GPS navigation apps (revisiting and updating the results from our previous work in Monogios et al. (2019)) and fitness tracking apps (which is a new study in relation with our work in Monogios et al. (2019)) on Android devices via performing dynamic analysis in order to identify which personal data – including user's device identifiers – they process. The dynamic analysis was carried out by using known appropriate software tools that help monitor what mobile applications are doing at runtime. We particularly put emphasis on whether such applications share the personal information they access with third-parties, due to the existence of third-party libraries. In the process, we also examined the privacy policies of these apps, in terms of finding out whether the information provided to the users is satisfactory. Our analysis shows that there is underlying data processing in place, which could possibly result in data protection risks, especially with respect to data leakage to third parties for tracking users, since the users are not fully aware of all these processes taking place at the background. Moreover, discrepancies occur with respect to the permissions that each application requires; again, since any such permission actually corresponds to a specific purpose of data processing, it seems that the relevant information provided to the users is not always adequate. Hence, this work further reveals the privacy challenges that span the entire mobile applications ecosystem.

It should be pointed out that the aim of the paper is not to make a comparative study between applications, neither to perform a legal analysis of the relevant personal data processing they perform; our aim is to examine, in a typical scenario of using any of these popular apps, which type of personal data processing occurs, so as to subsequently yield useful information on potential data protection concerns that are in place. Moreover, it should be also stressed that our analysis is based on specific software tools, as described next, which have some limitations in examining obfuscated and/or encrypted data and, therefore, it should not be seen as a complete analysis that suffices to cover any underlying personal data processing; in any case though, our findings allow for deriving safe conclusions.

The paper is organised as follows. First, a short presentation of previous relevant work is given in Section 2. Next, a discussion of the main legal provisions is given in Section 3, in conjunction with the presentation of device identifiers that should be considered as personal data. Section 4 provides a short overview on the permission model that Android adopts (since the permissions for accessing device information actually coincides with permissions for processing personal data), focusing on the so-called high-risk permissions in terms of privacy. Sections 5 and 6 constitute the main part of this work, where the results of our dynamic analysis on the corresponding applications are presented. More precisely, we first describe our testing environment and the methodology that have been utilised for our dynamic analysis, whilst we subsequently present the findings of the analysis, as well as a discussion on them. Finally, conclusion as well as some recommendations, are given in Section 7.

2 Related work

Privacy issues in the smart mobile ecosystem have been extensively studied in the literature, as also described in the Introduction. The extent to which users of smartphones can be uniquely identified simply through their personalised device configurations is studied in Kurtz et al. (2016). The aforementioned issue of the intra-library collusion is described in Taylor et al. (2017), illustrating how easily a third party can get access to (possibly) large volumes of user's personal data when the user simply installs a smart app and allows access to it through granting the relevant permissions asked. In Son et al. (2016), it is shown how malicious ads can infer sensitive information about users by accessing external storage; to this end, it is important to point out that even the mere existence of a specific type of file may reveal critical or even sensitive personal information (without needing to have read access to it). More recently, a comprehensive study on the privacy and security concerns associated with pre-installed Android software is presented in Gamba et al. (2020), illustrating – amongst others – that almost all such apps that have been identified as able to access personal data appear to disseminate that data to third parties; this is of high importance, taking into account that pre-installed apps typically run with privileged system permissions and, possibly, without the option of uninstalling.

Privacy issues have been also examined for specific-type smart applications, which due to their nature set additional privacy risks. In this regard, a comprehensive study of mobile health applications is provided in Papageorgiou et al. (2018), illustrating via both static and dynamic analysis that the majority of the analysed applications do not follow well-known practices. Moreover, a study on fitness and health apps is presented in Privacy Rights Clearinghouse (2013) which also raises several concerns; however, it is not clear, from the outcome of this report, which are the exact types of personal data that are being processed by third parties (as it is stated in Privacy Rights Clearinghouse (2013), almost all applications collect and send non-personally identifiable usage data to third parties for analysis, but it is not clear whether such data are indeed anonymous data). A very nice comparative study on fitness tracking apps, in terms of security and privacy features, is also given in Hilts et al. (2016). However, although this study comes to a conclusion – amongst other – that there is no transparency on the third parties that are getting access to users' data through the fitness trackers apps, an exact identification of such third parties, in the framework of intra-library collusion, is missing.

Moreover, the privacy concerns that are related with the general geotracking of an individual are well-known. For example, in a newspaper (see Valentino-Devries et al. (2018)) is described how the reporters were able to track an individual and learn a large amount of personal details just by examining the location data gathered by her smartphone. Such information is, for example, highly valuable for advertisers. However, more important consequences may also yielded. For example, a fitness-tracking app posted a heat map of its users across the world, having highlighted routes of the users; however, this heat map revealed the activities of US soldiers (see Hern (2018)). In a recent study (see Claesson et al. (2020)), the mobile advertising ecosystem has been particularly analysed, including the processing of GPS information; it is shown that several popular applications (not focusing explicitly on GPS navigating services) share the user's GPS location with multiple parties. The European Data Protection Board issued in 2020 guidelines on the use of location data and contact tracing tools in the context of the COVID-19 outbreak, emphasising that the contract tracing applications must not collect location data for the purpose of contact tracing (see EDPB (2020)).

Apart from the personal data processing performed by the smart apps (and the embedded libraries), privacy issues also arise from the mere information of which apps have been installed into a device. This is a direct consequence from the analysis in Tu et al. (2018) and Zhao et al. (2020), which illustrate how a predictive model can work well in deriving a profile of a user based simply on her/his applications usage. In our analysis we did not examine this privacy aspect; we focused on examining which types of personal data (including device data) are being accessed by these types of applications (GPS navigation apps and fitness trackers) as well as by the embedded third party libraries, whilst we also study how this underlying processing is transparent to the users.

3 Preliminaries: the notion of personal data protection

The European Union (2016) – known as the *General Data Protection Regulation* or GDPR – that applies from May 25th, 2018, constitutes the main legal instrument for personal data protection in Europe. The GDPR, which has been adopted in 2016 replacing the previous Data Protection Directive 95/46/EC, results in a harmonisation of relevant data processing rules across the European Union and aims to further protect and empower all EU citizens data privacy. Although the GDPR is a European Regulation, its territorial is not restricted within the European boundaries, since it applies to all organisations that process personal data of individuals residing in the European Union, regardless of the organisations' location, which can be outside European Union. As it is stated in Kaminski (2020), the intentionally global reach of the GDPR, in conjunction with the relevant threat of huge fines if fundamental rights are not properly protected, has led companies around the world to adjust their privacy practices – and countries around the world to update their privacy laws.

The term *personal data* refers to any information relating to an identified or identifiable natural person, that is a person who can be identified; as it is explicitly stated in the GDPR, an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. *Personal data processing* means any operation that is performed on personal data, including the collection, recording, structuring, storage, adaptation or alteration, retrieval, use, disclosure by transmission, dissemination, combination and erasure. The entity that, alone or jointly with others, determines the purposes and means of the processing of personal data, is the so-called *data controller*, whereas the entity which processes personal data on behalf of the controller is the *data processor*.

The notion of the personal data is quite wide, since special attention needs to be given whenever some data are being characterised as *anonymous*, i.e., non–personal. Indeed, according to the GDPR, although the data should be considered as anonymous if the person is no longer identifiable, all the means reasonably likely to be used to identify the natural person directly or indirectly should be taken into account towards determining whether a natural person is identifiable (see also Chatzistefanou et al. (2019)).

In general, device identifiers should be considered as personal data since they may allow the identification of a user (if possibly combined with other information). The Android operating system, which is the case considered in this work, is associated with two identifiers (see, e.g., Son et al. (2016)):

- the Android ID, which is a permanent 64bit randomly generated number
- the Google Advertising ID (GAID), which is a 32-digit alphanumeric identifier that can be reset at any time, according to the user's request.

Other device or network identifiers, such as the *medium access control* (MAC) and the *Internet protocol* (IP) addresses, should also be considered as personal data.

Some types of personal data are being mentioned as special categories of personal data; there are personal data related to racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, as well as data concerning health or data concerning a natural person's sex life or sexual orientation. Moreover, the processing of genetic data, as well as of biometric data for the purpose of uniquely identifying a natural person also correspond to processing of special categories of data. In general, there are stricter requirements for legitimate processing of such personal data (which are also being referred as sensitive data). It becomes evident that, depending on the scope and purpose of the smart application, it may be possible that it processes one or more special categories of personal data.

The GDPR codifies the basic principles that need to be guaranteed when personal data are collected or further processed and sets specific obligations to those that process personal data (data controllers/data processors). Any processing of personal data requires a lawful basis. In case that such a lawful basis is the individual's consent, then consent must meet certain requirements in order to be considered as being sufficient; more precisely, consent means any freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of his or her personal data must be given by a statement or a clear affirmative action (art. 4 of the GDPR). As stated in ENISA (2017), since many smart apps will need to rely on users' consent for the processing of certain personal data, the requirement of consent deserves special attention, in particular as it relates to the issue of *permissions*. Unfortunately, users have limited understanding of the associated risks of enabling permissions (or access to) in certain apps, whilst app developers have difficulties in comprehending and appropriately handling permissions (see ENISA (2017)). Moreover, this permissions model does not facilitate the provision of a legally valid consent for any third-party functionality that might be integrated into the app (since, in the Android platforms, third-party libraries inherit the privileges of the host app); hence, a major data protection risk occurs whenever a third-party library uses personal data for profiling and targeting, without the user's consent.

It should be pointed out that, depending on the techniques used, tracking of a mobile user may fall into the scope of the legal framework of the so-called cookie provision in the ePrivacy Directive (Directive 2002/58/EC); this applies only to the European Union. Again, this cookie provision requires informed consent for such app behaviour. In any case, the new Regulation that is currently being prepared in the EU to replace the ePrivacy Directive (the so-called ePrivacy Regulation), aims at being aligned with the GDPR, also covering new stakeholders and technologies in the field of electronic communications.

The GDPR sets new rules and obligations for any data controller. Amongst them, the so-called *data protection by design* and *data protection by default* constitute important challenges involving various technological and organisational aspects (see Alshammari et al. (2017)). According to the Recital 78 of the GDPR, appropriate measures that meet in particular the above two principles of data protection by design/default *(...) could consist, inter alia, of minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features.*

In the same Recital, there is also an explicit reference to the producers of the products, services and applications that are based on the processing of personal data or process personal data; namely, these stakeholders *(...) should be encouraged to take into account the right to data protection when developing and designing such products, services and applications and (...) to make sure that controllers and processors are able to fulfill their data protection obligations.* This is the only reference within the GDPR to stakeholders others than the data controllers or data processors. In the mobile ecosystem, application developers or library providers could lie in this category and thus, even in cases that these actors are neither data controllers nor data processors (hence, they may not be directly regulated under the GDPR), they are encouraged to make sure that controllers and processors are able to fulfill their data protection obligations (ENISA, 2017).

4 Permissions of applications

One of the Android system's core features is that applications are executed in their own private environment, referred to as a *sandbox*, being unable to access resources or perform operations outside of their sandbox that would adversely impact the system's security (e.g., by downloading malicious software) or user's privacy (e.g., by reading contacts, emails, or any other personal information) (see, e.g., Grammatikakis et al. (2018)). An application must explicitly request the permissions needed either at install-time, via its `AndroidManifest.xml` file, or at run-time. Our experimental environments, as it is discussed next, involved Android version 8.0 (API level 26), Android version 7.0 (API level 24) as well as Android Lollipop 5.0.1 (API level 21); therefore, for the first case the permissions granted to the applications were requested at runtime, whilst for the second and third case they were requested at install-time.

The permissions granted to applications are classified to several protection levels, based on their ability to harm the system or the end-user, out which three levels affect third-party applications (see Android Developers (2020)):

- *Normal permissions*: these cover areas where the application needs to access data or resources outside its sandbox, but where there's low risk to the user's privacy or the operation of other applications.
- *Signature permissions*: these are granted only if the application that requests the permission is signed by the same certificate as the application that defines the permission.
- *Dangerous permissions*: these cover areas where the application wants data or resources that involve the user's private information, or could potentially affect the user's stored data or the operation of other applications.

The permissions are in general strongly related with the notion of the user's consent for her personal data processing. However, in practice, the user is somehow forced to grant permissions, since otherwise the app cannot operate properly (and thus, her consent is actually not freely given). In addition, it is not always clear if the permissions granted are indeed strictly necessary for the functionality that the user asks, since the apps may require more permissions than actually needed to functioning properly. Bearing in mind the aforementioned issue of third parties that obtain the same privileges as the host app, it becomes evident that the permission model does not suffice to allow users having a sole control of her data processing, under full transparency. Besides, permissions are not a one-to-one mapping with the actual methods exposed by the API to manage the permissions – e.g., access to the camera, may also grant access to the photos automatically (see, e.g., ENISA (2017)).

5 Examination of GPS applications

This section provides the methodology that was followed, along with the results that have been obtained from the dynamic analysis performed on five popular GPS applications of the Android platform, which are available through the Google Play Store, namely:

1 the Google Maps (v. 10.12.1)

2 the Sygic GPS Navigation & Maps (v. 17.7.0)

3 the TomTom GPS Navigation – Traffic Alerts & Maps (v. 1.17.1)

4 the MAPS.ME (v. 9.0.7) and [5] the MapFactor GPS Navigation Maps (v. 4.0.109).

5.1 The testing environment

For our research experiments, we utilised an Android device (Android version Oreo 8) in which we installed the above five GPS navigation apps.

To be able to analyse these smart apps, via investigating whether they send personal data to third parties, as well as to obtain a direct information on potentially privacy–intrusive processes, we utilised the Lumen Privacy Monitor (Lumen),[1] which is a free, privacy–enhancing app with the ability to analyse network traffic on mobile devices in user space. The Lumen runs locally on the device and intercepts all network traffic since it inserts itself as a middleware between apps and the network interface (Razaghpanah et al., 2018). Lumen is able to identify personal data leaks that do not require explicit Android permissions, including software and hardware identifiers. Therefore, Lumen has been used in several cases by the research community for analysing potential personal data leakages from Android devices (see, e.g., Reyes et al. (2017)).

It should be noted that according to a communication we had with the team developing Lumen, it is possible that some leaks in Android 8.0 may not be detectable, since several apps use obfuscation or encoding to upload the data, even for location, and not all such mechanisms are supported in the public version of the Lumen. Therefore, we additionally performed an analysis through an appropriate module of the Xposed framework,[2] namely the Inspeckage Android packet inspector – that is an application with an internal HTTP server,[3] which is useful for performing dynamic analysis of Android applications. Due to practical limitations (i.e., we did not manage to operate smoothly this packet inspector in higher versions of Android), the Inspeckage Android packet inspector has been installed into a different device with an older version of the Android system, namely Android Lollipop 5.0.1; it should be pointed out though that, as of July 2019, the Lollipop versions had still about 14.5% share combined of all Android devices accessing Google Play store.[4] Since the same GPS applications, with the same embedded libraries, have been installed in both devices, it is expected that, for both scenarios we investigated, the same third-party domains collect data (differences may occur in which personal data the applications get access; for example, Android 8 does not allow applications getting access to the unique Android ID).

All the experiments took place during February and March 2019 in real (and not virtual) devices and, therefore, we were based on devices that we already possessed. The default settings were accepted during the installation of all GPS applications, whereas any permission that was required during their execution was also given, without examining whether each such permission is indeed absolutely necessary for the requested service.

5.2 Permission analysis of GPS applications

By using the Lumen tool, we observed the permissions that each of the application granted. We noticed that all applications asked for several access rights that are generally considered by the Lumen tool as high or medium risk with respect to user's privacy, such as the access to external storage and to the existing accounts on the device; all the permissions that are characterised as high-risk by Lumen are also considered as dangerous in Android Developers (2020). We summarise our observations, focusing explicitly on the so-called high-risk permissions, in Table 1.

Table 1 Dangerous permissions (`android.permission.⋆`) obtained by GPS navigation apps

Permissions	*[1]*	*[2]*	*[3]*	*[4]*	*[5]*
ACCESS_COARSE_LOCATION	×	×		×	×
ACCESS_FINE_LOCATION	×	×	×	×	×
READ_EXTERNAL_STORAGE	×	×	×	×	×
WRITE_EXTERNAL_STORAGE	×	×	×	×	×
CAMERA	×	×			
GET_ACCOUNTS		×	×		
RECORD_AUDIO	×	×			
READ_CONTACTS	×	×	×		
WRITE_CONTACTS			×		
READ_PHONE_STATE			×		×

It is of interest that, although all the applications provide similar services, there exist variations on the permissions that each of them requires. Therefore, the intra-library collusion privacy threat seems to be present; for example, if the same third-party library is being used by apps [1] and [3] or [2] and [3], then such a library will obtain all high-risk permissions that are shown in Table 1.

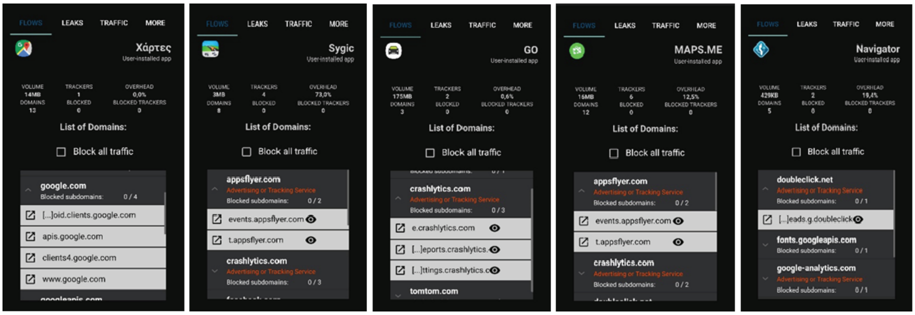
It should be explicitly pointed out that none of these permissions should be considered, by default, as unnecessary; for example, obviously, having location permission is prerequisite for GPS apps. Moreover, depending on the services provided, several other permissions may still be needed. However, it is questionable whether sufficient information is provided to the users regarding the necessity of these permissions, as well as whether third-party domains also get such permissions and have access to device data, as discussed next.

5.3 Data traffic to third-party domains by the GPS applications

By using the Lumen monitoring tool, we noticed that, for all GPS applications studied, there exists data traffic to several domains. With respect to Advertising Tracking Services (ATS), there exists – based on Lumen's output – one ATS in app [1], four ATS in app [2], two ATS in app [3], six ATS in app [4] and two ATS in app [5]. Indicative screenshots from the Lumen tool are provided in Figure 1.

By combining the outputs derived from both the Lumen and the Inspeckage tools, we summarise the results regarding the data leakages to several domains (either ATS or not) in Table 2; note that both first-party and third-party domains are shown. Based on these outputs, we conclude that, in most cases, the ATS that are associated with the apps are more than their number that was initially estimated by the Lumen tool.

Figure 1 Data leakages to several domains for the Google Map app, the Sygic app, the TomTom app, the MAPS.ME app and the Map Factor app respectively (see online version for colours)



FLOWS LEAKS TRAFFIC MORE

Χάρτες
User-installed app

VOLUME 14MB DOMAINS 13 TRACKERS 1 BLOCKED 0 OVERHEAD 0,0% BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

google.com
Blocked subdomains: 0 / 4
[...]oid.clients.google.com
apis.google.com
clients4.google.com
www.google.com

FLOWS LEAKS TRAFFIC MORE

Sygic
User-installed app

VOLUME 3MB DOMAINS 8 TRACKERS 4 BLOCKED 0 OVERHEAD 73,0% BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

appsflyer.com
Advertising or Tracking Service
Blocked subdomains: 0 / 2
events.appsflyer.com
t.appsflyer.com
crashlytics.com
Advertising or Tracking Service
Blocked subdomains: 0 / 3

FLOWS LEAKS TRAFFIC MORE

GO
User-installed app

VOLUME 175MB DOMAINS 3 TRACKERS 2 BLOCKED 0 OVERHEAD 0,6% BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

crashlytics.com
Advertising or Tracking Service
Blocked subdomains: 0 / 3
e.crashlytics.com
[...]eports.crashlytics.
[...]ttings.crashlytics.c
tomtom.com

FLOWS LEAKS TRAFFIC MORE

MAPS.ME
User-installed app

VOLUME 16MB DOMAINS 12 TRACKERS 6 BLOCKED 0 OVERHEAD 12,5% BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

appsflyer.com
Advertising or Tracking Service
Blocked subdomains: 0 / 2
events.appsflyer.com
t.appsflyer.com
crashlytics.com
Advertising or Tracking Service
Blocked subdomains: 0 / 2

FLOWS LEAKS TRAFFIC MORE

Navigator
User-installed app

VOLUME 429KB DOMAINS 5 TRACKERS 2 BLOCKED 0 OVERHEAD 19,4% BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

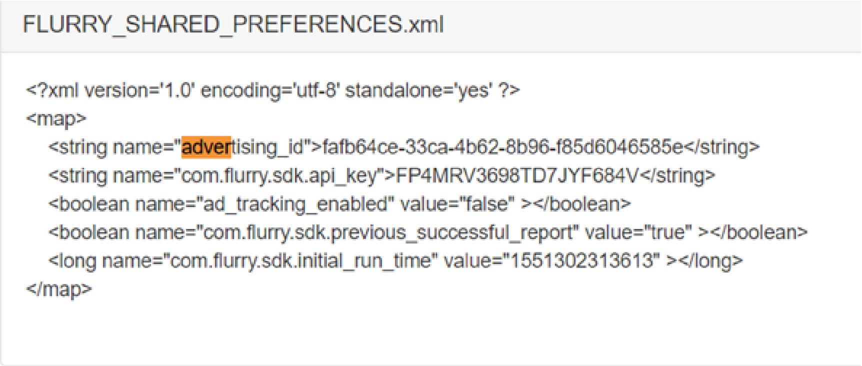
doubleclick.net
Advertising or Tracking Service
Blocked subdomains: 0 / 1
[...]eads.g.doubleclick
fonts.googleapis.com
Blocked subdomains: 0 / 1
google-analytics.com
Advertising or Tracking Service
Blocked subdomains: 0 / 1

Table 2 Data leakages by GPS navigation apps to several domains (either first or third-party)

Domains	*[1]*	*[2]*	*[3]*	*[4]*	*[5]*
app-measument.com	×				
google.com	×				×
youtube.com	×				
appsflyer.com		×		×	
crashlytics.com		×	×	×	
facebook.com		×		×	
foursquare.com		×			
infinario.com		×			
sygic.com		×			
uber.com		×			
windows.net		×			×
adjust.com			×		
tomtom.com			×		
flurry.com				×	
maps.me				×	
mopub.com				×	
my.com				×	
pushwoosh.com				×	
mapswithme.com				×	
mapfactor.com					×
google-analytics.com	×	×			×
googlesyndication.com					×
googleadservices.com					×
akamaized.net		×			
twitter.com		×	×	×	
doubleclick.net				×	×

We subsequently focused on the exact personal data, including device data, that are being transmitted to these domains. As explained previously, we utilised both the Lumen monitoring tool (for an Android 8) and the Inspeckage tool (for an Android Lollipop device). It should be pointed out that transmission of the GAID to third-party domains has been captured only by the Inspeckage tool, due to the encryption that takes place on such transmissions. An indicative screenshot on the information obtained by the Inspeckage tool is shown in Figure 2.

Figure 2 Transmission of the GAID to the flurry.com, based on the analysis through the Inspeckage tool (see online version for colours)



FLURRY_SHARED_PREFERENCES.xml

```
<?xml version='1.0' encoding='utf-8' standalone='yes' ?>
<map>
  <string name="advertising_id">fafb64ce-33ca-4b62-8b96-f85d6046585e</string>
  <string name="com.flurry.sdk.api_key">FP4MRV3698TD7JYF684V</string>
  <boolean name="ad_tracking_enabled" value="false" ></boolean>
  <boolean name="com.flurry.sdk.previous_successful_report" value="true" ></boolean>
  <long name="com.flurry.sdk.initial_run_time" value="1551302313613" ></long>
</map>
```

Our analysis illustrated that the GAID, as a unique device identifier, is being collected by several ATS services – namely by infinario.com (via app [2]), by appsflyer.com (via both apps [2] and [4]), by twitter.com (via [2], [3] and [4] apps), by flurry.com (via app [4]), by windows.net (via apps [2] and [5]) and by crashlytics.com (via app [4]).

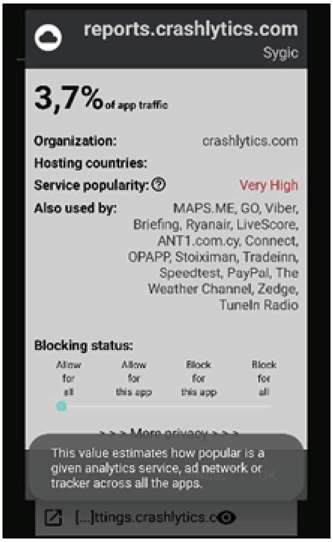
Interestingly enough, we noticed that there exist domains which may collect a combination of personal data due to the fact that are being embedded into several different apps. For example, the domain crashlytics.com collects the Facebook ID via the app [2]. Hence, if both apps [2] and [4] are being installed into the same device, both the GAID and the Facebook ID are being transmitted to this domain, thus allowing this ATS service linking a device with a social network user. Of course, it is also possible that such a pair – i.e., GAID and Facebook ID – are also being sent to an ATS service through a unique app; this is the case, e.g., of app [4] that sends these data to appsflyer.com. Moreover, it is highly probable that these domains may also collect user's information through other smart apps that are installed into the device, thus further increasing the privacy risks. For example, again for the Crachlytics tracking service, the Lumen tools informs us that several apps that are installed in our device also communicate with this domain; this is shown in Figure 3.

5.4 Transparency of the processing

With regard to the transparency of the underlying data processing, we studied the privacy policies of the GPS applications. It is interesting to point out that in most cases the privacy policies have been changed since the period in which the experiments took place, providing a more detailed information on the personal data that are being processed, as well the purposes of these processes (based on the checking of the privacy policies performed in August 2020). However, there may be still room for improvement in some cases. For example, in one case, the privacy policy states: *The Application uses GPS technology (or other similar technology) to determine your current location and display it on a map or during a turn-by-turn navigation. Your location history may be shared anonymously with third party partners (...) The Application uses Google Analytics for collecting anonymous usage information.*

This tracking information lets us better understand how you use the Application, allowing us to improve the user experience and correct any errors. No personally identifying data is included. You can disable analytics in Application settings (...). Such an information does not seem to be fully clear, taking into account that any device or network identifier should be considered as personal data and not as anonymous information.

Figure 3 The percentage of Sygic's outgoing traffic corresponding to the crashlytics service, as well as an enumeration of other apps in our device communicating with this service (including the two other GPS apps [2] and [3]) (see online version for colours)



reports.crashlytics.com
Sygic

3,7% of app traffic

Organization: crashlytics.com
Hosting countries:
Service popularity: Very High
Also used by: MAPS.ME, GO, Viber, Briefing, Ryanair, LiveScore, ANT1.com.cy, Connect, OPAPP, Stoiximan, Tradeinn, Speedtest, PayPal, The Weather Channel, Zedge, TuneIn Radio

Blocking status: [✓ Allow for all] Allow for this app Block for this app Block for all

> > > More privacy > > >

This value estimates how popular is a given analytics service, ad network or tracker across all the apps.

6 Examination of fitness tracking applications

This section provides the methodology that was followed, along with the results that have been obtained from the dynamic analysis performed on popular fitness tracking applications for the Android platform, which are available through the Google Play Store, namely: [F1] the Google Fit, [F2] the Samsung Health, [F3] the Mi Fit, [F4] the Huawei Health and [F5] the Garmin Connect. Each of them was communicating, through Bluetooth, with a wearable.

6.1 The testing environment

Again, as in the case of the GPS navigation apps, the Lumen Privacy monitor has been used to analyse the outgoing traffic from the above applications. However, to overcome limitations existing in the public version of the Lumen tool, as described previously, we proceeded with some modifications in our testing environment. First, we utilised a smart device with Android v. 7.0. Second, towards being able to decrypt the encrypted outgoing data from these applications, we used a laptop with the software OWASP Zed Attack Proxy (ZAP) v2.7, as well as the Burp Suite Community Edition v1.7.36, having the appropriate certificates installed in order to serve as an HTTPS proxy – i.e., to examine the outgoing from the Android device encrypted files. In other words, a somehow scenario of a Man-In-The-Middle (M.I.T.M.) attack was implemented, similarly to the case of an attacker aiming to read the encrypted data that are being sent by our device, without being detected (see, e.g., Alonso-Parrizas (2015)). However, as it will be shown next, there exist cases which the encrypted/obfuscated data could not be read, despite the fact that the relevant data flows were captured by the software tools used.

Each application has been examined separately each time – i.e., only one application was active each time. The applications were examined, in realistic conditions, for a period of 6 months (Apr. 2019 – Oct. 2019). As in the case of the GPS navigation apps, we used real (i.e., not virtual) devices that we possessed, whereas any permission that was required during their execution was given, without examining whether each such permission is indeed absolutely necessary for the requested service.

6.2 Permission analysis of fitness tracking applications

Similarly to the case of the GPS applications, we utilised the Lumen tool to observe the permissions that each of the fitness tracking application granted. Again, we summarise our observations, focusing explicitly on the so-called high-risk permissions, in Table 3; we omitted the reference to `BODY_SENSORS` permission, which is granted whenever the application communicates with a wearable.

Table 3 Dangerous permissions (`android.permission.`$\star$) obtained by fitness tracking apps

Permissions	$[F1]$	$[F2]$	$[F3]$	$[F4]$	$[F5]$
ACCESS_COARSE_LOCATION	×		×	×	×
ACCESS_FINE_LOCATION	×	×	×	×	×
READ_EXTERNAL_STORAGE	×	×	×	×	×
WRITE_EXTERNAL_STORAGE	×	×	×	×	×
CAMERA		×	×	×	×
GET_ACCOUNTS		×	×	×	×
RECORD_AUDIO			×	×	
READ_CONTACTS		×	×	×	×
READ_PHONE_STATE			×	×	×
ANSWER_PHONE_CALLS			×	×	×
READ_CALENDAR					×
SEND_SMS					×
CALL_PHONE			×	×	×
READ_CALL_LOG			×	×	×
READ_PHONE_NUMBERS				×	
PROCESS_OUTGOING_CALLS				×	

As in the case of GPS navigators, we observe that there exist variations on the permissions that each of them requires. Moreover, it is of interest to compare both Tables 1 and 3, in the context of intra-library collusion; if a third-party library is being used by a GPS application from Table 1 as well as by a fitness tracking application from Table 3, then the corresponding library provider obtains all permissions which are given to at least one of these host applications.

6.3 Data traffic to third-party domains

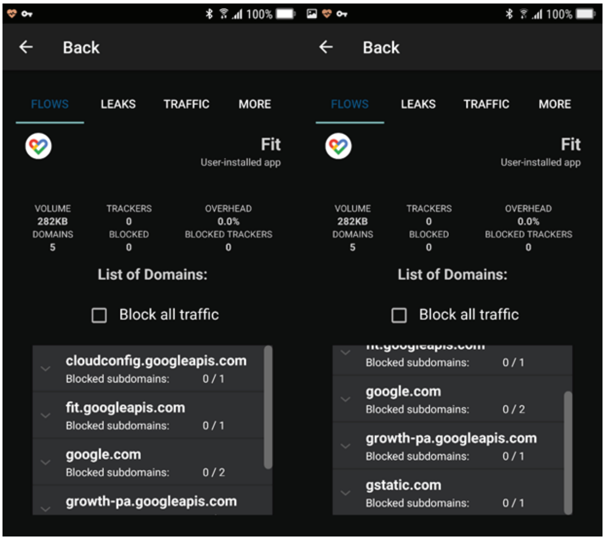
In this section, we provide detailed information on the outgoing – from the user's device – personal data induced by each of the above fitness tracking applications. For each case, the results are summarised, based on all software tools that have been used to capture and analyse this traffic.

6.3.1 Data traffic from fitness tracking app [F1]

The app [F1] uses sensors built into the device or watch to automatically track activities like walking, biking and running. The user may keep track of fitness goals that she sets, for specific time periods. The app has all the main features of fitness tracking applications – i.e., measuring steps, heart rate, performing automatic detection of activity such as walking, running, cycling, swimming, etc. It can be installed on all smart phone devices but also on wearable Android Wear devices.

According to the analysis through Lumen, the app [F1] sends data to domains belonging to Google and not to other (third-party) domains (see Figure 4). More detailed analysis on the type of data that are being transmitted by the application can be performed by utilising the OWASP ZAP tool. As shown in Figure 5, the email address of the user, the country (Cyprus) and the Android ID of the device are being transmitted to the app's servers, in conjunction with other device information such as the build fingerprint. The fact that the Android ID is being transmitted rests with the fact that our device uses the Android 7.0 version since, as it is already stated above, smart applications do not have access to this ID in subsequent Android versions.

Figure 4 Data leakage from app [F1] to other domains, based on analysis through Lumen (see online version for colours)



Fit
User-installed app

VOLUME 282KB
TRACKERS 0
OVERHEAD 0.0%
DOMAINS 5
BLOCKED 0
BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

cloudconfig.googleapis.com
Blocked subdomains: 0 / 1

fit.googleapis.com
Blocked subdomains: 0 / 1

google.com
Blocked subdomains: 0 / 2

growth-pa.googleapis.com

Fit
User-installed app

VOLUME 282KB
TRACKERS 0
OVERHEAD 0.0%
DOMAINS 5
BLOCKED 0
BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

Blocked subdomains: 0 / 1

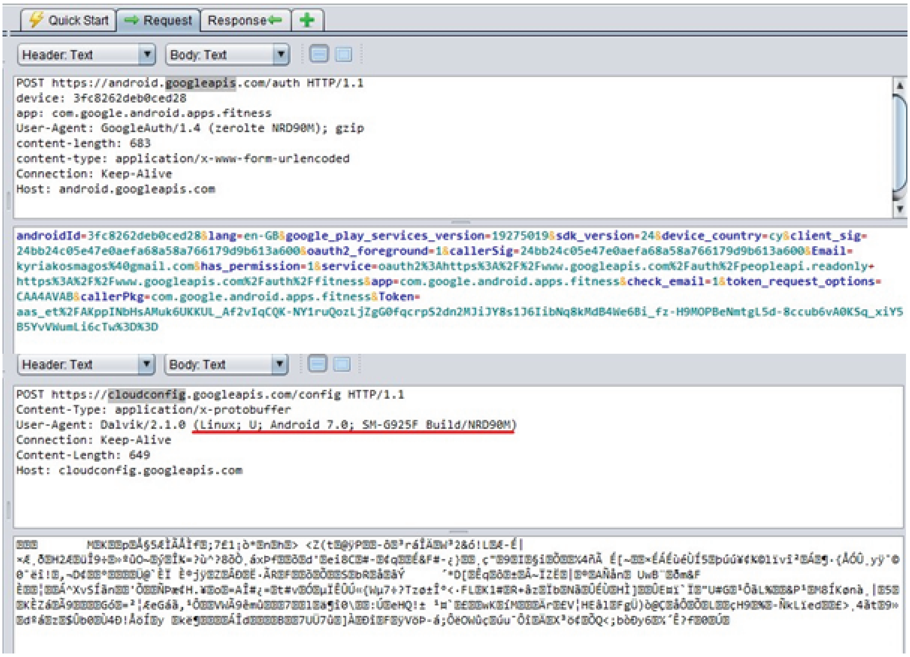
google.com
Blocked subdomains: 0 / 2

growth-pa.googleapis.com
Blocked subdomains: 0 / 1

gstatic.com
Blocked subdomains: 0 / 1

As any smart app provided by Google, there is a unified privacy policy covering any personal data processing by the company. Regarding the Google Fit app, we read at the app store (last accessed: June 2020) that the permissions for all versions of the app include storage (read the contents of the USB storage, modify or delete the contents of the USB storage), location (precise location and approximate location), wearable sensors/activity data (body sensors like heart rate monitors), identity (find accounts on the device, add or remove accounts), photos/media/files (read the contents of the USB storage, modify or delete the contents of the USB storage), contacts (find accounts on the device). It is not explicitly stated which of these permissions are strictly necessary and for what purpose.

Figure 5 Personal data that are being processed by the app [F1], based on analysis through ZAP (see online version for colours)



```
POST https://android.googleapis.com/auth HTTP/1.1
device: 3fc8262deb0ced28
app: com.google.android.apps.fitness
User-Agent: GoogleAuth/1.4 (zerolte NRD90M); gzip
content-length: 683
content-type: application/x-www-form-urlencoded
Connection: Keep-Alive
Host: android.googleapis.com
```

```
androidId=3fc8262deb0ced28&lang=en-GB&google_play_services_version=19275019&sdk_version=24&device_country=cy&client_sig=
24bb24c05e47e0aefa68a58a766179d9b613a600&oauth2_foreground=1&callerSig=24bb24c05e47e0aefa68a58a766179d9b613a600&Email=
kyriakosmagos%40gmail.com&has_permission=1&service=oauth2%3Ahttps%3A%2F%2Fwww.googleapis.com%2Fauth%2Fpeopleapi.readonly+
https%3A%2F%2Fwww.googleapis.com%2Fauth%2Ffitness&app=com.google.android.apps.fitness&check_email=1&token_request_options=
CAA4AVAB&callerPkg=com.google.android.apps.fitness&Token=
aas_et%2FAKppINbHsAMuk6UKKUL_Af2vIqCQK-NY1ruQozLjZgG0fqcrpS2dn2MJiJY8s1J6IibNq8kMdB4We6Bi_fz-H9MOPBeNmtgL5d-8ccub6vA0K5q_xiY5
B5YvVWumLi6cTw%3D%3D
```

```
POST https://cloudconfig.googleapis.com/config HTTP/1.1
Content-Type: application/x-protobuffer
User-Agent: Dalvik/2.1.0 (Linux; U; Android 7.0; SM-G925F Build/NRD90M)
Connection: Keep-Alive
Content-Length: 649
Host: cloudconfig.googleapis.com
```

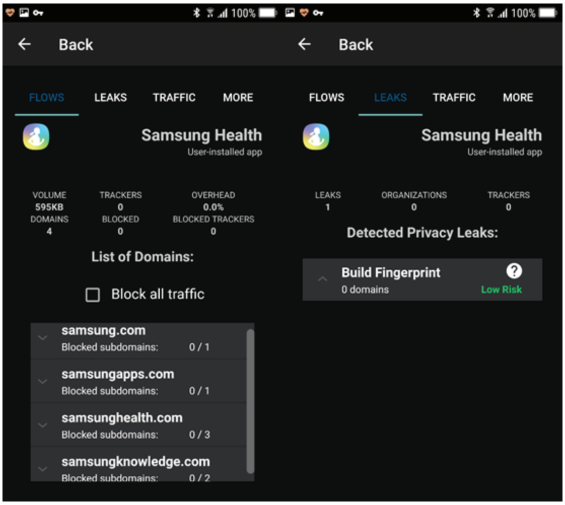
```
[illegible]
```

6.3.2 *Data traffic from fitness tracking app [F2]*

The app [F2] is an application tracking various aspects of daily life such as physical activity, diet, and sleep. Its main features include pedometer, calories monitoring, weight tracking, sleep monitoring etc. This application uses the sensors of the smartphone for data collection. In addition to the usual capabilities of these applications, the app can record water consumed, caffeine consumption, blood sugar etc.

According to the analysis through Lumen, the app [F2] sends data to domains belonging to the company and not to other (third-party) domains (see Figure 6). However, according to the analysis through the OWASP ZAP, it seems that an encrypted traffic is directed towards the domain clients4.google.com, but we did not manage to identify its content (see Figure 7). Moreover, by using the same tool, we noticed that information on user's lunch (which the user enters into the application) is being transferred towards a domain of another company which provides services related with calories counting (see Figure 8).

Figure 6 Data leakage from app [F2] to other domains, based on analysis through Lumen (see online version for colours)



Samsung Health
User-installed app

VOLUME 595KB
TRACKERS 0
OVERHEAD 0.0%
DOMAINS 4
BLOCKED 0
BLOCKED TRACKERS 0

List of Domains:

samsung.com
Blocked subdomains: 0 / 1

samsungapps.com
Blocked subdomains: 0 / 1

samsunghealth.com
Blocked subdomains: 0 / 3

samsungknowledge.com
Blocked subdomains: 0 / 2

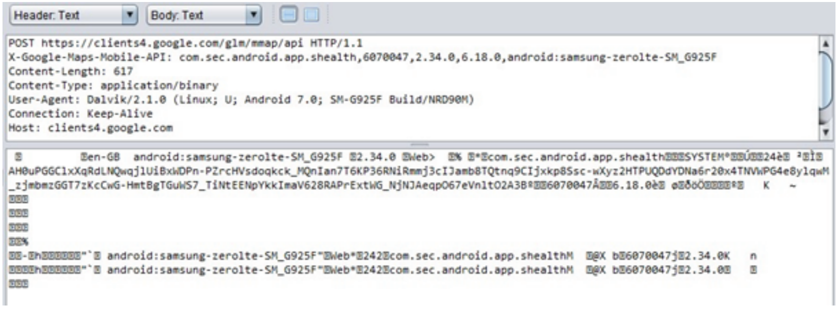
Samsung Health
User-installed app

LEAKS 1
ORGANIZATIONS 0
TRACKERS 0

Detected Privacy Leaks:

Build Fingerprint
0 domains
Low Risk

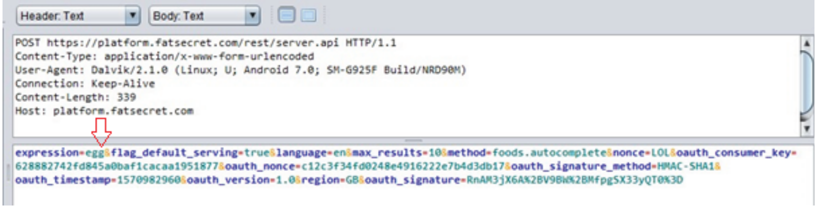
Figure 7 Encrypted/obfuscated data leakage from app [F2] to a google domain, based on the analysis through OWASP ZAP (see online version for colours)



```
POST https://clients4.google.com/glm/mmap/api HTTP/1.1
X-Google-Maps-Mobile-API: com.sec.android.app.shealth,6070047,2.34.0,6.18.0,android:samsung-zerolte-SM_G925F
Content-Length: 617
Content-Type: application/binary
User-Agent: Dalvik/2.1.0 (Linux; U; Android 7.0; SM-G925F Build/NRD90M)
Connection: Keep-Alive
Host: clients4.google.com
```

```
 ▯         ▯en-GB  android:samsung-zerolte-SM_G925F ▯2.34.0 ▯Web>  ▯% ▯*▯com.sec.android.app.shealth▯▯▯SYSTEM°▯▯Ú▯▯24è▯ ²▯Ì▯
AH0uPGGClxXqRdLNQwqjlUiBxWDPn-PZrcHVsdoqkck_MQnIan7T6KP36RNiRmmj3cIJamb8TQtnq9CIjxkp8Ssc-wXyz2HTPUQDdYDNa6r20x4TNVWPG4e8ylqwM
_zjmbmzGGT7zKcCwG-HmtBgTGuWS7_TiNtEENpYkkImaV628RAPrExtWG_NjNJAeqpO67eVnltO2A3B°▯▯6070047Å▯▯6.18.0è▯ ø▯ŏöÖ▯▯▯▯°▯   K   ~
▯▯▯
▯▯▯
▯▯▯
▯▯%
▯▯-▯h▯▯▯▯▯▯"`▯ android:samsung-zerolte-SM_G925F"▯Web*▯242▯com.sec.android.app.shealthM  ▯@X b▯6070047j▯2.34.0K   n
▯▯▯▯h▯▯▯▯▯▯"`▯ android:samsung-zerolte-SM_G925F"▯Web*▯242▯com.sec.android.app.shealthM  ▯@X b▯6070047j▯2.34.0▯   ▯
▯▯▯
```

Figure 8 Data traffic from app [F2] to another domain for calories counting, based on the analysis through OWASP ZAP (see online version for colours)



```
POST https://platform.fatsecret.com/rest/server.api HTTP/1.1
Content-Type: application/x-www-form-urlencoded
User-Agent: Dalvik/2.1.0 (Linux; U; Android 7.0; SM-G925F Build/NRD90M)
Connection: Keep-Alive
Content-Length: 339
Host: platform.fatsecret.com
```

```
expression=egg&flag_default_serving=true&language=en&max_results=10&method=foods.autocomplete&nonce=LOL&oauth_consumer_key=
628882742fd845a0baf1cacaa1951877&oauth_nonce=c12c3f34fd0248e4916222e7b4d3db17&oauth_signature_method=HMAC-SHA1&
oauth_timestamp=1570982960&oauth_version=1.0&region=GB&oauth_signature=RnAM3jX6A%2BV9BW%2BMfpgSX33yQT0%3D
```

The relevant page of the application in the app store (last accessed: August 2020) states that the app has access to (amongst others) precise location, storage (read, modify or delete of the

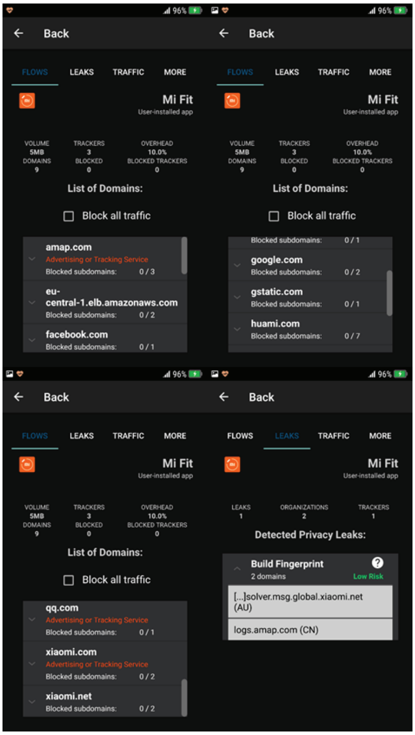
USB storage), identity and contacts (finding accounts in the device), camera and device ID. It is also stated in the privacy policy that data may also transmitted to third parties services, like Google Analytics. No explicit information on other third parties was given, whereas it is stated that the information transmitted to third parties include the device manufacturer, model and identifiers, as well as application identifiers. It is not mentioned though that such possibility of transmission is by default enabled.

6.3.3 Data traffic from fitness tracking app [F3]

The app [F3] is a popular fitness tracking app, which does not use the mobile phone as a device to collect data, but synchronises with other wearable company's devices and collects, analyses and processes data from them. Similarly to other fitness tracking apps, the [F3] is able to track user's activity, analyse sleep and evaluates the user's workouts.

According to the analysis through Lumen, the app [F3] sends data to nine (9) domains, whilst three of them are being characterised as ATS services – namely, the amap.com, the cgi.connect.qq.com and the xiaomi.net; other third parties domains though seem also to be included in the list of outgoing traffic (see Figure 9).

Figure 9 Data leakage from app [F3] to other domains, based on analysis through Lumen. (see online version for colours)



VOLUME 5MB
DOMAINS 9
TRACKERS 3
BLOCKED 0
OVERHEAD 10.0%
BLOCKED TRACKERS 0

List of Domains:

amap.com
Advertising or Tracking Service
Blocked subdomains: 0 / 3

eu-central-1.elb.amazonaws.com
Blocked subdomains: 0 / 2

facebook.com
Blocked subdomains: 0 / 1

Blocked subdomains: 0 / 1

google.com
Blocked subdomains: 0 / 2

gstatic.com
Blocked subdomains: 0 / 1

huami.com
Blocked subdomains: 0 / 7

qq.com
Advertising or Tracking Service
Blocked subdomains: 0 / 1

xiaomi.com
Advertising or Tracking Service
Blocked subdomains: 0 / 2

xiaomi.net
Blocked subdomains: 0 / 2

LEAKS 1
ORGANIZATIONS 2
TRACKERS 1

Detected Privacy Leaks:

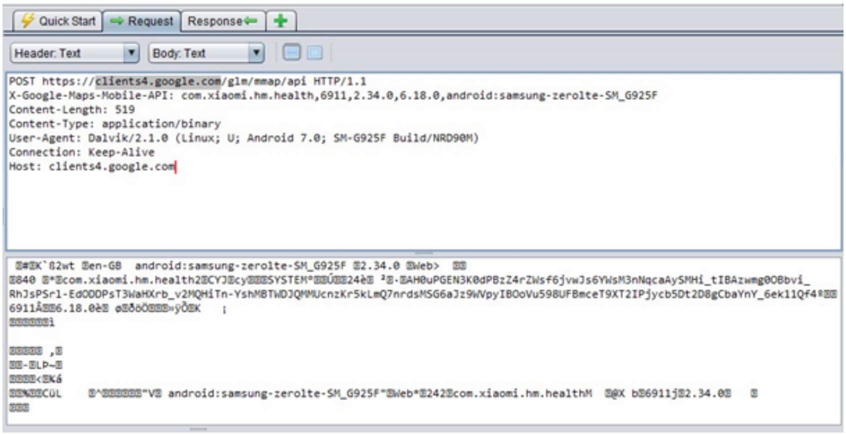
Build Fingerprint
2 domains
Low Risk

[...]solver.msg.global.xiaomi.net (AU)

logs.amap.com (CN)

The analysis due the OWASP ZAP confirmed the analysis of Lumen, illustrating data flows towards the app's servers (which include personal data such as device ID, name of user, birth data, sex, height, weight, as well as the data being collected by the wearable such as the heart rate) and third parties. However, similarly to the case of app [F2], the analysis through the OWASP ZAP indicated an encrypted/obfuscated data flow towards the domain clients4.google.com (see Figure 10) – it should be pointed out though that the Lumen had already detected the domain google.com as a third party domain receiving data. Moreover, the data being sent to the ATS amap.com are also encrypted and could not be identified (see Figure 11).

Figure 10 Encrypted/obfuscated data leakage from app [F3] to a google domain, based on the analysis through OWASP ZAP (see online version for colours)

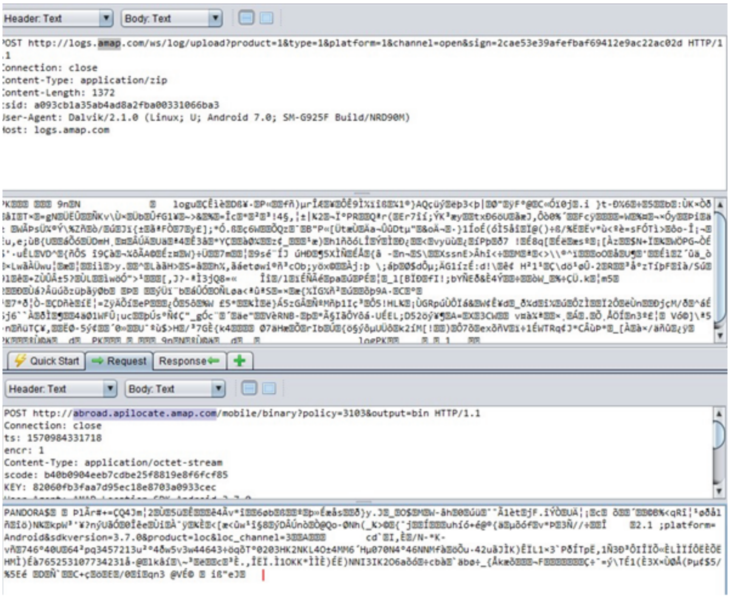


```
POST https://clients4.google.com/glm/mmap/api HTTP/1.1
X-Google-Maps-Mobile-API: com.xiaomi.hm.health,6911,2.34.0,6.18.0,android:samsung-zerolte-SM_G925F
Content-Length: 519
Content-Type: application/binary
User-Agent: Dalvik/2.1.0 (Linux; U; Android 7.0; SM-G925F Build/NRD90M)
Connection: Keep-Alive
Host: clients4.google.com
```

```
 □#□K`ß2wt □en-GB  android:samsung-zerolte-SM_G925F □2.34.0 □Web>  □□
□840 □*□com.xiaomi.hm.health2□CYJ□cy□□□SYSTEM°□□Ú□□24è□ ²□·□AH0uPGEN3K0dPBzZ4rZWsf6jvwJs6YWsM3nNqcaAySMHi_tIBAzwmg0OBbvi_
RhJsPSrl-EdODDPsT3WaHXrb_v2MQHiTn-YshMBTWDJQMMUcnzKr5kLmQ7nrdsMSG6aJz9WVpyIBOoVu598UFBmceT9XT2IPjycb5Dt2D8gCbaYnY_6ek11QF4®□□
6911Å□□6.18.0è□ ø□ðöÖ□□□»ÿÕ□K    ¡
□□□□□□i

□□□□□ ,□
□□-□LÞ~□
□□□□<□Ká
□□%□□CüL    □^□□□□□□"V□ android:samsung-zerolte-SM_G925F"□Web*□242□com.xiaomi.hm.healthM  □@X b□6911j□2.34.0□   □
□□□
```

Figure 11 Encrypted/obfuscated data leakage from app [F3] to the amap.com, based on the analysis through OWASP ZAP (see online version for colours)



```
'OST http://logs.amap.com/ws/log/upload?product=1&type=1&platform=1&channel=open&sign=2cae53e39afefbaf69412e9ac22ac02d HTTP/1
.1
:onnection: close
:ontent-Type: application/zip
:ontent-Length: 1372
:sid: a093cb1a35ab4ad8a2fba00331066ba3
Jser-Agent: Dalvik/2.1.0 (Linux; U; Android 7.0; SM-G925F Build/NRD90M)
Host: logs.amap.com
```

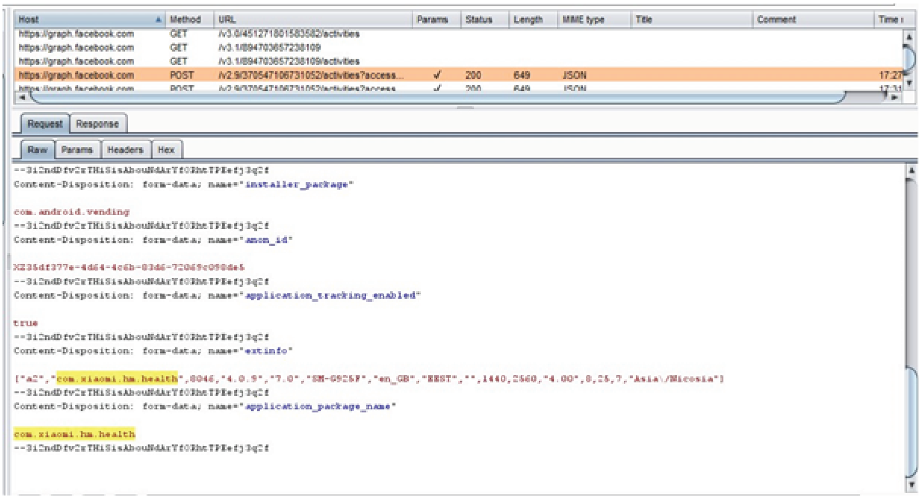
```
'K□□□ □□□ 9n□N            □   logu□ÇÊiè□D8¥·□P«□□fñ)µrÎÆ□¥□Ôê9Ì%ïß□%1®}AQçüýÿ□ep3<þ|□Ø"□ÿF°@□C«Óï0j□.i }t-Ð%6□+□5□□b□:ÙK×Òð
□àI□T×□=gN□ÜÊÛ□□ÑKv\Ù×□Üb□ÛfG1¥□~>&□%□=ÎC□*□²□³!4§,¦±|K2□¬Ï°PR□□Q*r(□Er7íí;ÝK³æy□□tx□6öU□åæJ,Ôò0%´□□FcÿÞ□□□=W□%¤□¬×ÓyÞ□PÍ□à
: □WÂÞsÜX°Ý\%ZÑ□ò/□ú□Jï{±□à*FÒ□7□y£];*Ó.ß□ç6W□□ÕQz□¨□B"P«[ÜtæÙ□Äa¬ÛûDtµ"□&oÄ¬□·}1ÍoÉ(óÌ5åî□Ï@()+ß/%Ë□Ev*ù<®è=sFÓTì>□ðo-Î¡¬□
:u,e;ùB{U□□áÕóÐÜDmH¸□¤□ÀÚÄ□U¤□*4□Ê3à□*YÇ□□à0%□□z¢_□□□¹æ)□h1ñõóLÎ□Ý□Ì□D¿□□<□vyüù□¿□íPþ□ð7 !□É8q[□Éê□es®□¡[Àz□□$N+Ï□K□WÖPG~ÒÉ
$"·uÊL□VD^□{ñÕS ï9Çà□¬%õÄA0□Ëz¤□W}+Ü□□7m□□¦□9sé¨ÍJ úHD□¶SXÌÑ□□ÉÅ□{å -□n¬□S\□□XssnE>Âhî<+□□M□*□<>\\®^ì□□□oO□å□U¶□'□□Éì□Z´ûä_ò
□×LwåÅÜwu¦□æ□¦□□ii□>y.□□^□LàãH>□S=å□□h%,äáetøwi®ñ³cOb;yöx©□□Àj:þ \;áþ□Ø$dÕµ;ÄG1ízÉ:d!\□é¢ H²1¹□Ç\dö¹øÛ·2□R□□³å°zTíþF□ìà/Sú□
)1□è□+ZÜÛÅ±5?□ÜL□□ìwöÓ">¹□□□[,J?-*Ì3jQ8=«     Ïï□/1□ïÉÑÃé□pa□ú□PÉ□¦□_1[BÏD□fI!;bÝÑEð&È4Ý□□+□□òW_□%+ÇÜ.k□¦m5□
:□□□Ð□Ùá?ÅüúðzüþäÿØb□ □Þ□ □□ýÜï¯b□áÛÓ□OÑLøa<*û*S□=×□æ{%ÏG%ñ²□Ú□□ðþ9A·□C□°□
'□7*ð¦Ò-□ÇDñèÈïÉ¦=ZÿÄÕí□ePÞ□□¿Ô□SÖ□%W £S*□□KÌ□è}ÁSzGÂ□Ñ*Mñþ1Iç³□ÕS!HLK□¦ÙGRpúÙÕÏá&□W¢Ê¥d□_ð%d□ì%□ú□ÕzÌ□□Ï2Õ□èÙn□□□jçM/ð□^áÉ
5j6``À□ðÌ□¶□□4ä01WFÛ¡uc□□þÚs°Ñ¢Ç"_gÓc¨□´□àe"□□VèRNB-□þ□*Ã§IäÕYöá·UÉEL;D52öÿ¥¶□A=□X□3CW□□ v¤à%*□□×¸□Á□.□Õ¸ÅÕÍ□n3*£¦□ Vó©]\*5
·n□ñüTÇ¥,□□ËØ-5ý¢□□´0«□□U¨*ù$>H□/³7GÈ{k4□□□□ Ø7äHæ□Õ□rIb□Ú□{ö§ýðµUÜð□k2íM[!□□)□Ô7ð□exðñV□ï+1ÉWTRq¢J*CÂùP*□_[À□à×/äñû□¿ÿ□
'K□□□□ÙØA□  d□  PK□□□ □ □□□ 9n□N□□ÙØA□  d□  □               lo□PK□□     □ □ 1   □□
```

```
POST http://abroad.apilocate.amap.com/mobile/binary?policy=3103&output=bin HTTP/1.1
Connection: close
ts: 1570984331718
encr: 1
Content-Type: application/octet-stream
scode: b40b0904eeb7cdbe25f8819e8f6fcf85
KEY: 82060fb3faa7d95ec18e8703a0933cec
```

```
PANDORA$□ □ Þ1År#+=ÇQ4Jm¦2□Ù□5ü□Ê□□□è4Åv*i□□6øb□ß□□*□þ»Éæås□□ð}y.J□_□O$□M□W-âh□0□ûÙ□¨"Ã1èt□jF.ïÝÒ□UÄ¦¡□c□ ð□□´□□©B%<qRî¦¹øðål
ñ□íö)NK□kpW³´¥?nýUãÓ□0ÎèeÙìIÊÀ¨ÿ□KÈ□<[æ<ûw¹ì§8□ýDÂÚnò□Ò@Qo-ØNh(_K>©□{¨j□□Í□□□uhíó+é@®{ä□µðóf□v*P□3Ñ//÷□□Î    □2.1 ;platform=
Android&sdkversion=3.7.0&product=loc&loc_channel=3□□A□□□       cd`□I,È□/N-*K-
vñ□746°40U□64²pq3457213u²°4ðw5v3w44643÷öqðT°0203HK2NKL4O±4MM6´Hµ070N4°46NNMfà□öÕu-42uãJÌK)ÊIL1×3`PðÍTpE,1Ñ3Ð³ÔIÎÏÕ«ÈLÌÏÍÓEÈÕE
HMÌ)Éà765253107734231å-@□1kåí□\~³□e□□c□³È.,ÎEÏ.Ì1OKK*ÌÎÈ)ÉE)NNI3IK2O6aðó□+cbà□`äbø+_{Åkæð□□□¬F□□□□□□□Ç+¯=ý\TÉ1(È3X×ÙØÅ(Þµ¢$5/
%5Eé □D□Ñ`□□C+ç□ö□E□/0□ï□qn3 @VÉ© □ iß"eJ□
```

In addition, the analysis through the Burp Suite shows that there is a data leakage towards a Facebook domain, which includes the GAID of the device – see Figure 12.

Figure 12 Data leakage from app [F3] to the Facebook, based on the analysis through the Burp Suite (see online version for colours)



Host	Method	URL	Params	Status	Length	MIME type	Title	Comment	Time
https://graph.facebook.com	GET	/v3.0/451271801583582/activities							
https://graph.facebook.com	GET	/v3.1/894703657238109							
https://graph.facebook.com	GET	/v3.1/894703657238109/activities							
https://graph.facebook.com	POST	/v2.9/370547106731052/activities?access...	✓	200	649	JSON			17:27
https://graph.facebook.com	POST	/v2.9/370547106731052/activities?access	✓	200	649	JSON			17:31

```
--3i2ndDfv2rTHiSisAbouNdArYf0RhcTPEefj3q2f
Content-Disposition: form-data; name="installer_package"

com.android.vending
--3i2ndDfv2rTHiSisAbouNdArYf0RhcTPEefj3q2f
Content-Disposition: form-data; name="anon_id"

XZ35df377e-4d64-4c6b-83d6-72069c098de5
--3i2ndDfv2rTHiSisAbouNdArYf0RhcTPEefj3q2f
Content-Disposition: form-data; name="application_tracking_enabled"

true
--3i2ndDfv2rTHiSisAbouNdArYf0RhcTPEefj3q2f
Content-Disposition: form-data; name="extinfo"

["a2","com.xiaomi.hm.health",8046,"4.0.9","7.0","SM-G925F","en_GB","EEST","",1440,2560,"4.00",0,25,7,"Asia\/Nicosia"]
--3i2ndDfv2rTHiSisAbouNdArYf0RhcTPEefj3q2f
Content-Disposition: form-data; name="application_package_name"

com.xiaomi.hm.health
--3i2ndDfv2rTHiSisAbouNdArYf0RhcTPEefj3q2f
```

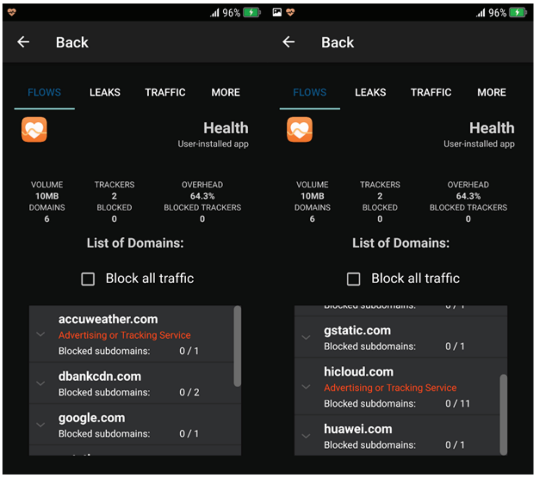
The relevant page of the application in the app store (last accessed: August 2020) states that the app has access to (amongst others) wifi connections, precise location, storage (read, modify or delete of the USB storage), identity and contacts(finding accounts in the device), camera, microphone (record audio), application history (retrieve running apps), call logs and device ID. It is also stated in the privacy policy that the company *may collect certain information automatically through its websites, products, services or other methods of analysis, such as your Internet protocol (IP) address, cookie identifiers, mobile carrier, mobile advertising identifiers, MAC address, and other device identifiers that are automatically assigned to your computer or device when you access the internet, browser type and language, geolocation information, hardware type, operating system, internet service provider, pages that you visit before and after using the services, the date and time of your visit, the amount of time you spend on each page, information about the links you click and pages you view within the services, and other actions taken through use of the services such as preferences*; moreover, regarding data transfers to third parties, it is explicitly mentioned that there exists data transmission for advertising and third party marketing, for which the user *has the option to opt-out*; this obviously implies – although it is not explicitly mentioned – that such data transfers are by default enabled.

6.3.4 Data traffic from fitness tracking app [F4]

The application [F4] has the same capabilities as [F3], supporting, e.g., walking, running, cycling mode, record running track, heart rate, trajectory and other sports data. The application, like [F3], is based on the same company's wearable devices for data collection, which are synchronised with it towards analysing and processing data. It should be mentioned here that if the application is needed to be installed into a device from another manufacturer, the user must install the relevant Mobile Services application from the Play Store.

According to the analysis through Lumen, the app [F4] sends data to six (6) domains, whilst two of them are being characterised as ATS services – namely, the accuweather.com and the hicloud.com; other third parties domains though seem also to be included in the list of outgoing traffic (see Figure 13). Again, as in the previous cases, the data flow towards the domain clients4.google.com was in an encrypted/obfuscated form, which could not be read (see Figure 14), whereas the Lumen monitor identified that the domain configdownload.dbankcdn.com collected the device fingerprint (this is being characterised as a privacy leak of low risk).

Figure 13 Data leakage from app [F4] to other domains, based on analysis through Lumen (see online version for colours)



Health
User-installed app

VOLUME 10MB
TRACKERS 2
OVERHEAD 64.3%
DOMAINS 6
BLOCKED 0
BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

accuweather.com
Advertising or Tracking Service
Blocked subdomains: 0 / 1

dbankcdn.com
Blocked subdomains: 0 / 2

google.com
Blocked subdomains: 0 / 1

Health
User-installed app

VOLUME 10MB
TRACKERS 2
OVERHEAD 64.3%
DOMAINS 6
BLOCKED 0
BLOCKED TRACKERS 0

List of Domains:

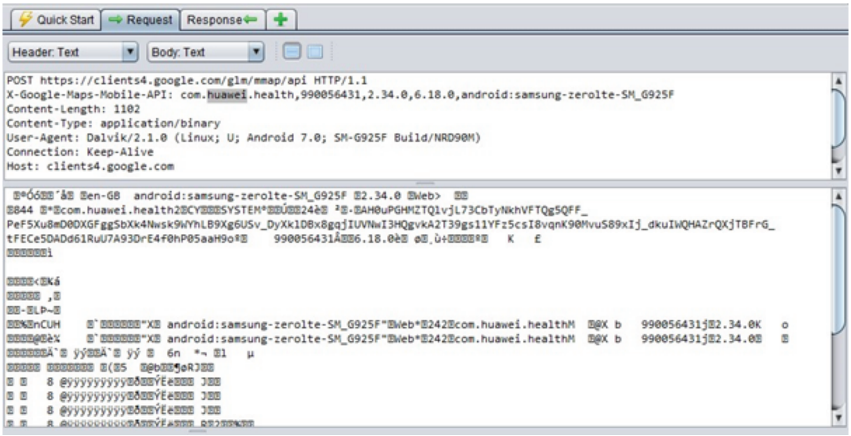
□ Block all traffic

gstatic.com
Blocked subdomains: 0 / 1

hicloud.com
Advertising or Tracking Service
Blocked subdomains: 0 / 11

huawei.com
Blocked subdomains: 0 / 1

Figure 14 Encrypted/obfuscated data leakage from app [F4] to a google domain, based on the analysis through OWASP ZAP (see online version for colours)



```
POST https://clients4.google.com/glm/mmap/api HTTP/1.1
X-Google-Maps-Mobile-API: com.huawei.health,990056431,2.34.0,6.18.0,android:samsung-zerolte-SM_G925F
Content-Length: 1102
Content-Type: application/binary
User-Agent: Dalvik/2.1.0 (Linux; U; Android 7.0; SM-G925F Build/NRD90M)
Connection: Keep-Alive
Host: clients4.google.com
```

```
 ▯ªÓó▯▯´å▯ ▯en-GB  android:samsung-zerolte-SM_G925F ▯2.34.0 ▯Web>  ▯▯
▯844 ▯*▯com.huawei.health2▯CY▯▯▯SYSTEMª▯▯Ú▯▯24è▯ ²▯-▯AH0uPGHMZTQ1vjL73CbTyNkhVFTQg5QFF_
PeF5Xu8mD0DXGFggSbXk4Nwsk9WYhLB9Xg6USv_DyXk1DBx8gqjIUVNwI3HQgvkA2T39gs11YFz5csI8vqnK90MvuS89xIj_dkuIWQHAZrQXjTBFrG_
tFECe5DADd61RuU7A93DrE4f0hP05aaH9oª▯    990056431Å▯▯6.18.0è▯ ø▯¸ù÷▯▯▯▯ª▯   K   £
▯▯▯▯▯▯ì

▯▯▯▯<▯Ká
▯▯▯▯▯ ,▯
▯▯-▯LÞ~▯
▯▯%▯nCUH    ▯`▯▯▯▯▯▯"X▯ android:samsung-zerolte-SM_G925F"▯Web*▯242▯com.huawei.healthM  ▯@X b   990056431j▯2.34.0K   o
▯▯▯▯@▯èX    ▯`▯▯▯▯▯▯"X▯ android:samsung-zerolte-SM_G925F"▯Web*▯242▯com.huawei.healthM  ▯@X b   990056431j▯2.34.0▯   ▯
▯▯▯▯▯▯Ä`▯ ÿý▯▯Ä`▯ ÿý ▯  6n  *¬ ▯1   µ
▯▯▯▯▯ ▯▯▯▯▯▯▯ ▯(▯5  ▯@b▯▯¶øRJ▯▯
▯ ▯   8 @ÿÿÿÿÿÿÿÿÿ▯ð▯▯ÝÉe▯▯▯ J▯▯
▯ ▯   8 @ÿÿÿÿÿÿÿÿÿ▯ð▯▯ÝÉe▯▯▯ J▯▯
▯ ▯   8 @ÿÿÿÿÿÿÿÿÿ▯ð▯▯ÝÉe▯▯▯ J▯▯
```

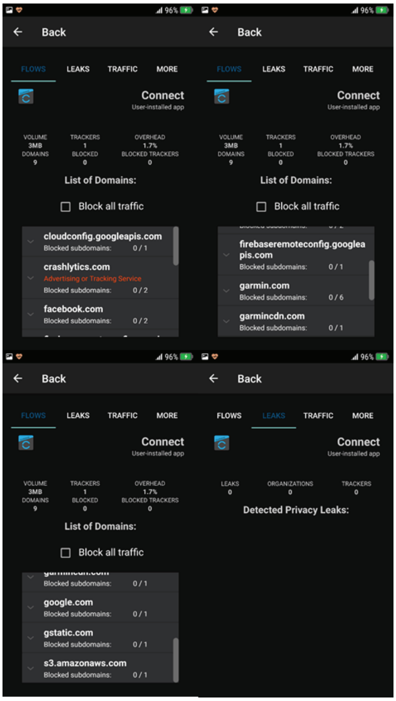
An interesting observation is that the other software tools in our testing environment did not manage to identify more concrete lists of personal data being transmitted to app's servers or third parties – and thus, we were not able to have a clear picture of the underlying personal data processing that occurs in real time.

The relevant page of the application in the app store (last accessed: August 2020) states that the app has access to (amongst others) precise location, storage (read, modify or delete of the USB storage), identity and contacts(finding accounts in the device), camera, microphone, WiFi connections, call logs and device ID. It is also stated in the privacy policy that if the user logs in with an account such as Google, Facebook, or Twitter, third-party account information such as account identifier, email address, nickname, profile picture and date of birth are being obtained. It is also mentioned that the purposes of the personal data processing include analytics and development purposes. It is stated that these processes are based on the legitimate interests of the company, which in turn implies that the user's consent is not required for such a processing (and, thus, we get that these processes are by default enabled).

6.3.5 Data traffic from fitness tracking app [F5]

The company supporting [F5] offers a wide range of smart watches, which can also be used for fitness tracking. The application [F5], similarly to the previous ones, is synchronised with these devices and collects and analyses the user's data, which subsequently, after processing them, presents them to her.

Figure 15 Data leakage from app [F5] to other domains, based on analysis through Lumen (see online version for colours)



Connect
User-installed app

VOLUME 3MB
TRACKERS 1
OVERHEAD 1.7%
DOMAINS 9
BLOCKED 0
BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

cloudconfig.googleapis.com
Blocked subdomains: 0 / 1

crashlytics.com
Advertising or Tracking Service
Blocked subdomains: 0 / 2

facebook.com
Blocked subdomains: 0 / 2

Connect
User-installed app

VOLUME 3MB
TRACKERS 1
OVERHEAD 1.7%
DOMAINS 9
BLOCKED 0
BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

firebaseremoteconfig.googleapis.com
Blocked subdomains: 0 / 1

garmin.com
Blocked subdomains: 0 / 6

garmincdn.com
Blocked subdomains: 0 / 1

Connect
User-installed app

VOLUME 3MB
TRACKERS 1
OVERHEAD 1.7%
DOMAINS 9
BLOCKED 0
BLOCKED TRACKERS 0

List of Domains:

□ Block all traffic

Blocked subdomains: 0 / 1

google.com
Blocked subdomains: 0 / 1

gstatic.com
Blocked subdomains: 0 / 1

s3.amazonaws.com
Blocked subdomains: 0 / 1

Connect
User-installed app

LEAKS 0
ORGANIZATIONS 0
TRACKERS 0

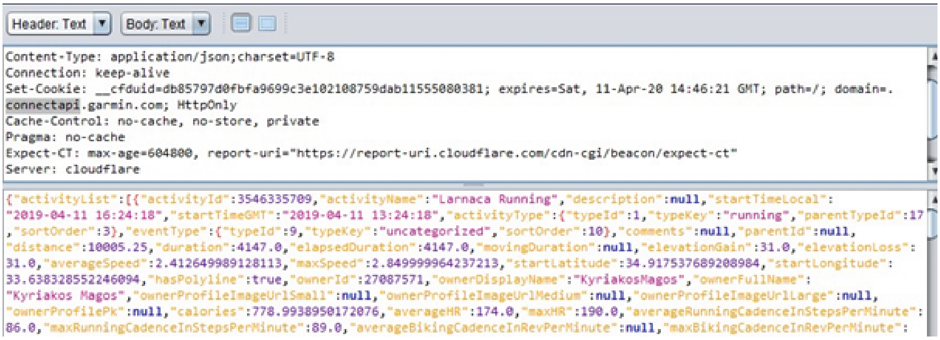
Detected Privacy Leaks:

According to the analysis through Lumen, the app [F5] sends data to nine (9) domains, whilst one of them is being characterised as ATS service – namely, the crashlytics.com (the one that was also included in the GPS applications discussed in Section 5.3); other third parties domains though seem also to be included in the list of outgoing traffic (see Figure 15). As in all the other fitness tracking applications that were examined, the data towards the Google domain could not be read in our testing environment.

The usage of the Burp Suite tool allowed for a more detailed information on the underlying personal data processing. More precisely, the personal data that are being transmitted to the app's servers (user name, number of calories, heart rate, type of activity etc.) are indicatively shown in Figure 16. In addition, the Burp suite tool revealed a data flow towards Facebook, which includes the GAID of the device (see Figure 17).

The relevant page of the application in the app store (last accessed: August 2020) states that the app has access to (amongst others) precise location, storage (read, modify or delete of the USB storage), identity and contacts(finding accounts in the device), camera, wifi connections, call logs, calendar and device ID. It is also stated in the privacy policy that the user's data are being used for provision of personalised stats and insights (such as how many calories the user burns during an activity and how the user compares to similar users), whereas if the user chooses to upload data from the Garmin device to Garmin Connect, the company collects recorded data from those activities so the user can analyse it and track the progress toward her goals. Such an uploading to the company's server is not mandatory; it is up to the user whether she desires to upload them or keep them locally on her device (in the latter case though, some features will not be supported). In addition, the privacy policy explicitly states that the company *will not share the user's activity data with third parties unless the user tells to.*

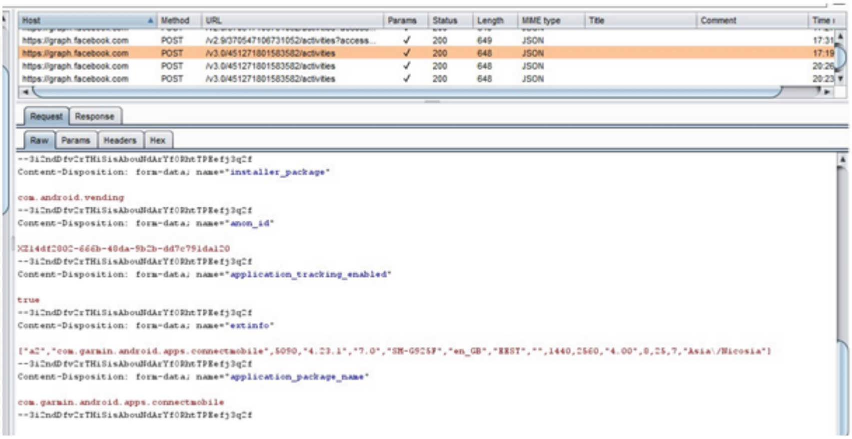
Figure 16 Data flow from app [F5] to the app's server, based on analysis through the Burp Suite (see online version for colours)



```
Content-Type: application/json;charset=UTF-8
Connection: keep-alive
Set-Cookie: __cfduid=db85797d0fbfa9699c3e102108759dab11555080381; expires=Sat, 11-Apr-20 14:46:21 GMT; path=/; domain=.
connectapi.garmin.com; HttpOnly
Cache-Control: no-cache, no-store, private
Pragma: no-cache
Expect-CT: max-age=604800, report-uri="https://report-uri.cloudflare.com/cdn-cgi/beacon/expect-ct"
Server: cloudflare
```

```
{"activityList":[{"activityId":3546335709,"activityName":"Larnaca Running","description":null,"startTimeLocal":
"2019-04-11 16:24:18","startTimeGMT":"2019-04-11 13:24:18","activityType":{"typeId":1,"typeKey":"running","parentTypeId":17
,"sortOrder":3},"eventType":{"typeId":9,"typeKey":"uncategorized","sortOrder":10},"comments":null,"parentId":null,
"distance":10005.25,"duration":4147.0,"elapsedDuration":4147.0,"movingDuration":null,"elevationGain":31.0,"elevationLoss":
31.0,"averageSpeed":2.412649989128113,"maxSpeed":2.849999964237213,"startLatitude":34.917537689208984,"startLongitude":
33.638328552246094,"hasPolyline":true,"ownerId":27087571,"ownerDisplayName":"KyriakosMagos","ownerFullName":
"Kyriakos Magos","ownerProfileImageUrlSmall":null,"ownerProfileImageUrlMedium":null,"ownerProfileImageUrlLarge":null,
"ownerProfilePk":null,"calories":778.9938950172076,"averageHR":174.0,"maxHR":190.0,"averageRunningCadenceInStepsPerMinute":
86.0,"maxRunningCadenceInStepsPerMinute":89.0,"averageBikingCadenceInRevPerMinute":null,"maxBikingCadenceInRevPerMinute":
```

Figure 17 Data flow from app [F5] to the Facebook, based on analysis through the Burp Suite (see online version for colours)



Host	Method	URL	Params	Status	Length	MIME type	Title	Comment	Time
https://graph.facebook.com	POST	/v2.9/370547106731052/activities?access...	✓	200	649	JSON			17:31
https://graph.facebook.com	POST	/v3.0/451271801583582/activities	✓	200	648	JSON			17:19
https://graph.facebook.com	POST	/v3.0/451271801583582/activities	✓	200	648	JSON			20:26
https://graph.facebook.com	POST	/v3.0/451271801583582/activities	✓	200	648	JSON			20:23

```
--3i2ndDfv2rTHiSisAbouNdArYf0RhtTPEefj3q2f
Content-Disposition: form-data; name="installer_package"

com.android.vending
--3i2ndDfv2rTHiSisAbouNdArYf0RhtTPEefj3q2f
Content-Disposition: form-data; name="anon_id"

XZ14df2802-666b-48da-9b2b-dd7c791da120
--3i2ndDfv2rTHiSisAbouNdArYf0RhtTPEefj3q2f
Content-Disposition: form-data; name="application_tracking_enabled"

true
--3i2ndDfv2rTHiSisAbouNdArYf0RhtTPEefj3q2f
Content-Disposition: form-data; name="extinfo"

["a2","com.garmin.android.apps.connectmobile",5090,"4.23.1","7.0","SM-G925F","en_GB","EEST","",1440,2560,"4.00",0,25,7,"Asia\/Nicosia"]
--3i2ndDfv2rTHiSisAbouNdArYf0RhtTPEefj3q2f
Content-Disposition: form-data; name="application_package_name"

com.garmin.android.apps.connectmobile
--3i2ndDfv2rTHiSisAbouNdArYf0RhtTPEefj3q2f
```

7 Conclusion

In this paper we studied some popular smart applications for Android platforms which process personal data such as geographical location and fitness/health data, with the aim to examine whether they suffer from known privacy issues that are present in the mobile applications ecosystem, taking into account relevant legal provisions. The main findings of our analysis can be summarised as follows:

- The information that is provided to the users regarding the relevant underlying personal data processing is not always complete or clear (although it seems that the corresponding privacy policies are constantly getting improved in terms of the information they provide). Moreover, in some cases there exist some data flaws to third parties for which there is no explicit information.
- It is possible that some applications, taking into account their access rights, process some personal data in a way that it is questionable if the data protection by design and by default principles are being met. For example, it is not always clear why some specific Android permissions are required by an application.
- The known privacy threat that rests with the so-called intra-library collusion seems to exist – since there exist common ATS services in more than one applications.

The above findings are consistent with the known data protection issues that are present in the mobile applications ecosystem. Although such issues are known to exist for several types of smart apps (see, e.g., Chatzistefanou et al. (2019), Icram and Kaafar (2017) and Ren et al. (2018)), the fact that they are also present in applications processing user's location and fitness/health has significant importance as it may result higher privacy violations.

It should be stressed that the above findings do not necessarily constitute an exhaustive list; since the main data flows are encrypted for security reasons, there exist significant restrictions on analysing potential data leakages and, thus, there is still room for further analysis of these apps. However, despite these restrictions, it becomes evident that much effort should be put on promoting the data protection by design and by default principles in smart applications such as privilege separation strategies for apps and their embedded libraries as well as improvement on personal data policies (both on their content/clarity but on their ease on readability). All the relevant stakeholders – namely app developers, library providers, operating system providers, app stores – have a crucial role on this and, thus, it is essential to build further upon their knowledge and awareness. To this end, some conclusions on possible recommendations can be stated as follows:

- Full transparency of the underlying personal data processing, including the cases of third parties, is needed. The users should explicitly know the embedded third-party tracking and advertising services and their providers, as well as the types of data they collect. Ambiguities regarding which data are anonymous and which are not should be eliminated in the relevant information provided.
- Moreover, full information of which data processing is by default enabled or not should be also provided, in relation with the corresponding permissions that the user grants.
- The above necessitates a new type of (more detailed) privacy policies, as well as embedded mechanisms to allow users provide their explicit informed consent. Both application developers and operating system providers should put effort on supporting such features, whereas the role of the app stores is also important towards ensuring that the apps satisfy these requirements.
- The notion of which is the strictly necessary personal data processing for each case should be further elaborated. Since this is strongly related with the so-called data protection by default principle, the application providers should be accountable on justifying the necessity of each of the underlying personal data processings that are considered as strictly necessary, whereas the relevant information provided should also clarify this aspect.

Finally, it is important to point out that regulators have also an important role in addressing the aforementioned issues. Taking into account that the European Union is in the process of updating the current e-Privacy Directive into a new e-Privacy Regulation, which is expected to apply to several data processes as those described in this paper (being aligned with the GDPR), it is essential to establish specific requirements on transparency and personal data protection by design and by default, covering all the above stakeholders.

Acknowledgement

The authors would like to thank the anonymous reviewers for their useful comments and suggestions, which helped to improve the paper.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 786698. The work reflects only the authors' view and the Agency is not responsible for any use that may be made of the information it contains.

References

Alonso-Parrizas, A. (2015) *Forensic Analysis on Android: A Practical Case*, SANS Institute, https://www.sans.org/reading-room/whitepapers/mobile/forensic-analysis-android-practical-case-36317.

Alshammari, M., Simpson, A. (2017) 'Towards a principled approach for engineering privacy by design', in Schweighofer, E., Leitold, H., Mitrakas, A. and Rannenberg, K. (Eds): *Privacy Technologies and Policy (APF) 2017*, LNCS, Springer, Heidelberg, Vol. 10518, pp.161–177.

Android Developers (2020) *Permissions Overview*, https://goo.gl/A7QG1J (Accessed 29 August, 2020).

Athanasopoulos E., Kemerlis V.P., Portokalidis G. and Keromytis A.D. (2016) 'NaClDroid: native code isolation for Android applications', in Askoxylakis, I., Ioannidis, S., Katsikas, S. and Meadows, C. (Eds.): *Computer Security – ESORICS 2016*, LNCS, Springer, Cham, Vol. 9878, pp.422–439.

Binns, R., Lyngs, U., Van Kleek, M., Zhao, J., Libert, T. and Shadbolt, N. (2018) *Third Party Tracking in the Mobile Ecosystem*, arXiv:1804.03603v3 [cs.CY].

Bujlow, T., Carela-Español, V., Solé-Pareta, J., Barlet-Ros, P. (2017) 'A survey on web tracking: mechanisms, implications, and defenses', *Proceedings of the IEEE*, Vol. 105, pp.1476–1510.

Castelluccia, C. (2012) 'Behavioural tracking on the internet: a technical perspective', in Gutwirth, S., Leenes, R., De Hert, P. and Poullet, Y. (Eds.): *European Data Protection: In Good Health?*, Springer, Heidelberg, pp.21–33.

Chatzistefanou, V., Limniotis, K. (2019) 'Anonymity in social networks: The case of anonymous social media', *International Journal of Electronic Governance (IJEG)*, Inderscience Publishers, Vol. 11, Nos. 3–4, pp.361–385.

Chester, J. and Montgomery, K.C. (2017)'The role of digital marketing in political campaings', *Internet Policy Review: Journal on internet regulation*, Vol. 6, pp.1–20.

Claesson, A. and Bjørstad, T.E. (2020) *Out of Control – A Review of Data Sharing by Popular Mobile Apps*, Technical Report, Norwegian Consumer Council (2020), https://fil.forbrukerradet.no/wp-content/uploads/2020/01/mnemonic-security-test-report-v1.0.pdf

Eckersley, P. (2010) 'How unique is your web browser?', *10th Int. Symp. on Private Enhancing Technologies (PETS)*, pp.1–18.

European Data Protection Board (2020) *Guidelines 04/2020 on the Use of Location Data and Contact Tracing Tools in the Context of the COVID-19 Outbreak*, https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_20200420_contact_tracing_covid_with_annex_en.pdf

European Union Agency for Cybersecurity (2017) 'Privacy and data protection in mobile applications – A study on the app development ecosystem and the technical implementation of GDPR'. Available: https://www.enisa.europa.eu/publications/privacy-and-data-protection-in-mobile-applications

European Union (2016) 'Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)', *Official Journal L.*, Vol. 119, No. 1, pp.1–88.

Gamba, J., Rashedy, M., Razaghpanahz, A., Tapiador, J. and Vallina-Rodriguez, N. (2020) 'An analysis of pre-installed Android software', *IEEE Symposium on Security and Privacy*, pp.1039–1055 (virtual event).

Gervais, A., Filios, A., Lenders, V., Capkun, S. (2017) 'Quantifying Web adblocker privacy', in Foley, S.N., Gollmann, D. and Snekkenes, E. (Eds.): *22nd European Symposium on Research in Computer Security (ESORICS), Part II*, LNCS, Springer, Vol. 10493, pp.21–42.

Hern, A. (2018) 'Fitness tracking app Strava gives away location of secret US army bases', *The Guardian*, https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases

Hilts, A., Parsons, C. and Knockel, J. (2016) *Every Step You Fake: A Comparative Analysis of Fitness Tracker Privacy and Security*, Open Effect Report, https://openeffect.ca/reports/Every_Step_You_Fake.pdf

Grammatikakis, K-P., Ioannou, A., Shiaeles, S. and Kolokotronis, N. (2018) 'Are cracked applications really free? An empirical analysis on Android devices', *16th IEEE International Conference on Dependable, Autonomic and Secure Computing (DASC)*, pp.730–735.

GSM Association (2017) *Safety, Privacy and Security Across the Mobile Ecosystem – Key Issues and Policy Implications*, https://www.gsma.com/publicpolicy/wp-content/uploads/2017/02/GSMA_Safety-privacy-and-security-across-the-mobile-ecosystem.pdf

Ikram, M. and Kaafar, M.A. (2017) 'A first look at mobile Ad-Blocking apps', *IEEE 16th Int. Symp. on Network Computing and Applications (NCA)*, pp.1–8.

Kaminski, M. (2020) 'A recent renaissance in privacy law', *Communications of the ACM*, Vol. 63, No. 9, pp.24–27.

Krumm, J. (2010) 'Ubiquitous advertising: The killer application for the 21st century', *IEEE Pervasive Computing*, Vol. 10, pp.66–73.

Kurtz, A., Gascon, H., Becker, T., Rieck, K. and Freiling. F. (2016) 'Fingerprinting mobile devices using personalized configurations', *Proceedings on Privacy Enhancing Technologies*, pp.4–19.

Mattke, J., Müller, L.K. and Maier, C. (2017) 'Why do individuals block online ads? An explorative study to explain the use of ad blockers', *23rd Americas Conference on Information Systems (AMCIS)*, Association for Information Systems.

Mavriki, P. and Karyda, M. (2019) 'Big data in political communication: implications for group privacy', *International Journal of Electronic Governance (IJEG)*, Vol. 11, Nos. 3–4, pp.289-309, Inderscience Publishers.

Mazel, J., Garnier, R. and Fukuda, K. (2017) *A Comparison of Web Privacy Protection Techniques*, arXiv:1712.06850v1 [cs.CR].

Monogios, S., Limniotis, K., Kolokotronis, N., Shiaeles, S (2019) 'A case study of intra-library privacy issues on Android GPS navigation apps', in Katsikas, S. and Zorkadis, V. (Eds.): *E-Democracy – Safeguarding Democracy and Human Rights in the Digital Age. e-Democracy 2019*, Communications in Computer and Information Science, Springer, Cham, Vol. 1111, pp.33–48.

Pang, H. (2018) 'Mobile communication and political participation: unravelling the effects of mobile phones on political expression and offline participation among young people', *International Journal of Electronic Governance (IJEG)*, Inderscience Publishers, Vol. 10, No. 1, pp.3–23.

Papageorgiou, A., Strigkos, M., Politou, E.A., Alepis, E., Solanas, A. and Patsakis, C. (2018) 'Security and privacy analysis of mobile health applications: the alarming state of practice', *IEEE Access*, Vol. 6, pp.9390–9403.

Pew Research Center (2013) *Anonymity, Privacy, and Security Online*, http://pewinternet.org/Reports/2013/Anonymity-online.aspx

Privacy Rights Clearinghouse (2013) *Mobile Health and Fitness Apps: What Are the Privacy Risks?*, Technical Report (2013), https://privacyrights.org/consumer-guides/mobile-health-and-fitness-apps-what-are-privacy-risks

Razaghpanah, A., Nithyanand, R., Vallina-Rodriguez, N., Sundaresan, S., Allman, M., Kreibich, C. and Gill, P. (2018) 'Apps, trackers, privacy, and regulators: a global study of the mobile Tracking ecosystem', *Network and Distributed System Security Symposium*, San Diego, California, USA.

Ren, J., Lindorfer, M., Dubois, D.J., Rao, A., Choffnes, D. and Vallina-Rodriguez, N. (2018) 'Bug fixes, improvements, ... and privacy leaks – a longitudinal study of PII leaks across Android App versions', *Network and Distributed System Security Symposium*, San Diego, California, USA.

Reyes, I., Wijesekera, P., Razaghpanah, A., Reardon, J., VallinaRodriguez, N., Egelman, S. and Kreibich, C. (2017) 'Is our children's apps learning? automatically detecting coppa violations', *IEEE Workshop on Technology and Consumer Protection (ConPro)*, San Jose, CA, USA.

Son, S., Kim, D. and Shmatikov, V. (2016) 'What mobile ads know about mobile users', *Network and Distributed System Security Symposium*, San Diego, California, USA.

Statista (2020) *Number of Smartphone Users Worldwide from 2016 to 2021*, https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

Statista (2019) *Share of Global Smartphone Shipments by Operating System from 2014 to 2023*, https://www.statista.com/statistics/272307/market-share-forecast-for-smartphone-operating-systems/

Stevens, R., Gibler, C., Crussell, J., Erickson, J., Chen, H. (2012) 'Investigating user privacy in Android Ad libraries', *Workshop on Mobile Security Technologies (MoST)*, p.10, San Fransisco, USA.

Taylor, V.F., Beresford, A.R. and Martinovic, I. (2017) *Intra-Library Collusion: A Potential Privacy Nightmare on Smartphones*, arXiv:1708.03520v1 [cs.CR] (2017).

Tu, Z., Li, R., Wang, G., Wu, D., Hui, P., Su, L. and Jin, D. (2018) 'Your apps give you away: distinguishing mobile users by their app usage fingerprints', *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, Vol. 2, No. 3, pp.138:1–138:3.

Valentino-Devries, J., Singer, N., Keller, M.H. and Krolik, A. (2018) 'Your apps know where you were last night, and they're not Keeping it secret', *NY Times article*, https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html

Zhao, S., Xu, Y., Ma, X., Jiang, Z., Luo, Z., Li, S., Yang, L. T., Dey, A.K. and Pan, G. (2020) 'Gender profiling from a single snapshot of apps installed on a smartphone: an empirical study', *IEEE Transactions on Industrial Informatics*, Vol. 16, No. 2, pp.1330–1342.

Notes

[1] https://www.haystack.mobi/
[2] https://repo.xposed.info
[3] https://mobilesecuritywiki.com/
[4] https://developer.android.com/about/dashboards