Android apps and advertising networks – a survey on data privacy

Dirk Pawlaszczyk*

Hochschule Mittweida – University of Applied Sciences,
Technikumplatz 17, Mittweida,
09648, Germany
Email: pawlaszc@hs-mittweida.de
*Corresponding author

Jannik Weber, Ralf Zimmermann and
Christian Hummert

Central Office for Information Technology in the
Security Sector (ZITiS),
Zamdorfer Str. 88, Munich,
81677, Germany
Email: jannik.weber@zitis.bund.de
Email: ralf.zimmermann@zitis.bund.de
Email: christian.hummert@zitis.bund.de

Abstract: Advertising networks bring companies together that run advertisements on websites or within apps. Targeting capabilities of the advertisers have increased drastically over the last years due to the availability user data and advances in data science: whether users log on to social networks or use a mobile app, they leave their traces everywhere and leave valuable data for analysis. However, even in the background and unconsciously to many users, personal data is stored and processed by third parties. This data is used to get to know potential customers better and to align the appropriate advertising with them automatically. This work analyses selected advertising networks in Android apps concerning their data collection behaviour. Therefore a selection of 100 apps from the Google Play Store has been analysed on their contained advertising networks. The network traffic originating from the given apps as well as the app behaviour has been analysed. It is pointed out which data is collected and transferred. The results are quite surprising stating that the top app comprising 14 advertising networks. The conformity and completeness of the information provided in the data privacy declarations will also be assessed.

Keywords: data privacy; advertising networks; Android apps; digital forensics.

Reference to this paper should be made as follows: Pawlaszczyk, D., Weber, J., Zimmermann, R. and Hummert, C. (2020) ‘Android apps and advertising networks – a survey on data privacy’, *Int. J. Information Privacy, Security and Integrity*, Vol. 4, No. 4, pp.261–275.

Biographical notes: Dirk Pawlaszczyk pursued a Diploma in Computer Science from the Technical University of Ilmenau in 2000 and a PhD of Science in 2009. He is currently working as a Full Professor in the Department of Computer Sciences, Hochschule Mittweida – University of Applied Sciences. He has published more than 20 research papers in reputed inter-national journals, including Springer and IEEE, and it is also available online. His main research work focuses on digital forensics, network security, cloud security and privacy, IoT, distributed simulation, and artificial intelligence. He has eight years of teaching experience and 12 years of research experience.

Jannik Weber has received his Bachelor in General and Digital Forensic Science from the Hochschule Mittweida – University of Applied Sciences. He is currently absolving his Master's in Cybersecurity at the Bundeswehr University Munich. His current area of research focus primary on cryptanalysis and reverse engineering.

Ralf Zimmermann graduated with a Diploma in Computer Science at the Technische Universität Braunschweig and started his research in the field of embedded security and applied cryptanalysis at the Ruhr-Universität Bochum. Finishing his PhD, he joined the Forensic Science Institute of the German Federal Criminal Police Office (BKA) and is currently the Head of Research in Cryptanalysis at the ZITiS. His research interests cover applied cryptanalysis, high-performance implementations, digital forensics and reverse engineering.

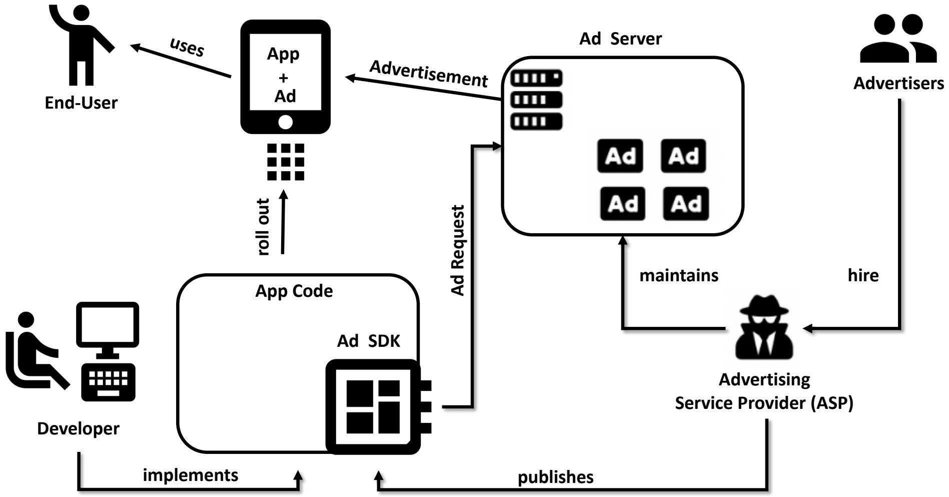
Christian Hummert is the Director of the Digital Forensics Section at the ZITiS. Before, he was a Full Professor for IT-Security/Digital Forensics at the Faculty of Applied Computer and Biosciences of Mittweida University of Applied Sciences. He has graduated in Computer Sciences at the Albert Einstein University in Ulm. After his Doctorate at the Friedrich Schiller University in Jena, he worked six years as a forensic expert for the Federal Police of Thuringia. He was appointed to the Interpol Digital Forensics Expert Group. Several research projects in his group are focusing on digital forensics. He is especially interested in malware analysis and car forensics. He is involved in the expert training for digital forensics at the German BKA.

1 Introduction

For the majority of the people today, the mobile phone is a constant companion in our everyday life: in the mobile age, we take one thing for granted – that we can access everything from everywhere. There seems to be a suitable app to help with almost every problem and usually these apps are easy to install and – in most cases – free of charge. However, many users forget that only few services are for free and that they pay a service- or use-fee with sensitive and private information about themselves. The business model of app developers has changed within the last years and money is no longer solely earned through software licenses: Advertisements placed in the applications are used to increase the profit. In order to provide the user with tailored advertisements, different sets of data are collected, often without the explicit consent (sometimes even without the knowledge of the user) of websites and apps.

So-called advertising networks have played a central role in this for many years. In order to understand the systematic approach, the reasoning behind and the implications of advertisements (ads) in apps and advertisement networks, we will first introduce the different stakeholders and continue with the necessary steps leading to the placement and display of an ad. Figure 1 shows the different players in the advertisement system.

Figure 1 An overview over the stakeholders in the advertisement business model and the actions necessary to roll out ads to the end-user



These include but are not limited to *advertisers*, *advertising service providers* (ASPs), *developers*, and *end-users* (Kees and Andrews, 2019; Vallina-Rodriguez et al., 2016):

- *ASPs* are companies, who consist of an advertising agency and also maintain a network of advertising servers. The advertising agency interacts with the customers and offers different marketing strategies (or promotion campaigns). In an advertising campaign, the budget and the target number of clicks are usually calculated for a specific period of time.

  Besides the maintenance of the advertising network and the corresponding servers to distribute ads through different channels, the ASP also develops and publishes software development kits (SDKs) for different platforms to access the network.

- *Advertisers* are the companies trying to improve the awareness about their products and services. For this purpose, the advertisers hire ASPs and provide the advertising motifs in form of texts, images, and videos. These materials are then uploaded to the ASP network and will be used according to the selected advertisement campaign.

- *Developers* design applications for one or multiple platforms. In this case, we focus on mobile devices. The developers use the provided SDKs of ASPs in their apps. Using ads allows them to provide the software for free (no purchase required) and still gain profit from the user-base.

- The *end-users group* is the last participant in this system, who desire to use (free) apps on their mobile devices. By downloading, installing and actively using the software, the advertisement SDK interacts with the advertising servers, downloading and presenting ads to the end-user.

To give support to advertisements within mobile applications, the ASPs provide their own advertisement SDKs, which the developers integrate into their apps (Son et al., 2016; Stevens et al., 2012). It includes the programming interface to access the network of the ASP, through which the developer can request advertisements to display in the application. The SDK performs an *ad request* to the *ad server* in the advertising network of the service provider. Executing such a request automatically attaches information about the user, which the ad server uses to select a suitable advertisement (Son et al., 2016). To realise this, the software automatically gathers the additional information the ASP expects, processes the returned data and displays the advertisement in the program (Narayanan et al., 2014). Accordingly, the developers do not need to implement these features on their own and – by using multiple SDKs – is able to query multiple ASPs without additional effort.

From the perspective of data privacy the practice of current advertising networks is at least questionable. The techniques used by the ASPs are referred to as *behavioural targeting* (Yuan et al., 2012; Federal Trade Commission, 2020). The data collected by the advertising networks are used to target online advertising to consumers based on preferences inferred from their online doings. European and American data protection authorities and data privacy activist consider the current practice of processing customer data extremely precarious (Esteve, 2017; Graef and Berlo, 2020). In their opinion, the ASP information strategies are not compliant with applicable data privacy regulations (CNIL, 2018; Data Protection Commission, 2019). In the past, politicians and data protection specialists have repeatedly criticised such networks for the way they handle customer data. In essence, the responsible authorities blame the assembling of information from different sources. They argue that customers are not sufficiently informed and that the disclosure of such data is subject to approval (Federal Trade Commission, 2016). Transparency about terms and conditions is not enough. Users are not informed 'at the time of installation' about the advertising purpose and the identity of the responsible data processors. The customers are not sufficiently informed about all transferred data. On the other hand, ASPs like Google, firmly committed to the privacy of their advertisers and users (Google Inc., 2020). However, it is problematic that the providers do not pass on the data protection agreements directly but always only indirectly via the terms and condition of the respective app. Which position corresponds to the truth can only be judged if it is clear which data are transmitted and how they are processed.

Before a statement can be made to this question, we have to make clear, which data is really transferred from the respective programs to the advertising network? In the context of this publication, we shed light on the question what data is sent by popular apps. Therefore, this paper discusses selected advertising networks in Android apps concerning their data collection behaviour. Therefore a selection of 100 apps from the Google Play Store have been analysed on their contained advertising networks. For each app, we examined two main points: how much personal information does the users usually share with the app and which smartphone options of the app are accessed? The conformity and completeness of the information provided in the data privacy declarations will also be assessed.

2 Related work

In the past years, advertising networks have attracted the attention of security researchers. In their research, they focused on different aspects of security and privacy threats. We will briefly summarise the related work.

Already in 2012, a software to analyse the security and privacy risks called AdRisk was published in (Grace et al., 2012). This framework was used to systematically identify potential risks in mobile apps. Its main goal was the automatic analysis of advertising SDKs, identifying potential risks the libraries impose on the mobile device. In this study, about 100,000 apps from Google Play Store were used to first identify the most commonly used advertising SDKs, which were analysed afterwards. The authors concluded that most libraries send private information – such as location information, caller lists, the telephone number, browser bookmarks, and the list of installed applications – to the ad network. Some libraries may even download additional software from the internet and execute the code on the smartphone, which imposes critical security risks.

Crussell et al. (2014) examine the problem of ad fraud, where malevolent code fetches advertisements in the background from ad servers without user interaction. Such fraud apps attempt to remain stealthy when generating ad traffic by only periodically sending clicks. The authors address a special form of abuse in advertisements. However, it does not reflect on data privacy.

Another contribution focuses on the effects of advertising SDKs on the privacy of an end-user (Stevens et al., 2012). They studied the privacy impact of 13 advertising SDKs by analysing the authorisations used. The authors showed considerable vulnerabilities related to privacy. They point out, that some advertising SDKs permissions were queried, without being documented beforehand. These authorisations were to read private data when available. These included permissions to access the camera function as well as to the calendar and contacts. Authorisations for advertising SDKs have also been analysed with regard to the evolution of those apps across multiple versions (Book et al., 2013). For this purpose, advertising SDKs from 114,000 apps has been examined. They were looking at the release dates of apps in which the software libraries are included. The result of this research was that the use of permissions had increased over the years – the use of permissions that pose a particular risk to privacy and user safety.

In another publication, Kim et al. (2018) argue for more transparency in the disclosure of personal information in apps. For this purpose, the researchers examine the influence of transparency and the effectiveness of advertising in different studies.

Another major problem in this context is the malicious use of SDKs and how to identify them. Backes et al. (2016) for example, address the question how to recognise SDKs of advertisers. The study concludes that advertising SDK brings additional security vulnerabilities to their host apps. Some of them also misuse the permissions indirectly inherited via the app.

Beyond this, different contributions deal with the automatic recognition of software libraries (Narayanan et al., 2014) or the identification and characterisation of domains related to advertisements and user tracking (Vallina-Rodriguez et al., 2016).

Advertising SDKs were also analysed in regard to how they protect the users of advertising-financed apps from malicious advertising (Son et al., 2016). The authors found that SDKs generally protect the user from malicious advertising by loading the advertisements into a protected environment within the application. The authors also showed how sensitive information from users could be protected against malicious advertisements while granting access to the memory of the application.

Finally, there is a series of preliminary work on the subject of target-specific advertising. Advertisers usually want to reach a specific target group with their campaigns. Accordingly, this type of advertising is called targeting (Dogruel, 2019). There is a number of parameters to widen the success of a campaign by concentrating on a particular target group. For example, advertisers can use the parameters for advertising products or services that are only relevant to a particular region (Lian et al., 2019). It is also possible to carry out advertising campaigns specifically for women or men, depending on gender. Advertisements can also be placed depending on the network connection in which the target person is located. That information is used, among other things, by the advertising providers to decide whether to place a video ad or a picture ad and thus to reduce mobile data consumption.

In this work, we build upon the basis created by the related work and focus on the data privacy aspects and possible issue in our analysis.

3 Analysis

This study targets the most commonly used advertising networks and the impact on data protection and privacy. Before we identify these networks or look at the application implementing advertisements, we briefly describe our test setup including the hardware and software we used to analyse the Android applications. As hardware platform, we acquired a consumer Android phone, which supported Android 9 (Pie) and was not difficult to root (in this case by using Magisk). We chose a *Xiaomi Mi A2* device, but any other smartphone with root support will suffice.

On the host side, we use a standard Linux system and control the smartphone via the Android Debug Bridge (ADB) (Google Developers, 2020), which provides an interface to perform device actions such as installing and debugging apps.

As the previous studies, we limited the number of apps to the top 100 most popular free applications. Google automatically defines categories and collections, which also include a collection of applications with respect to the number of downloads/installations. We use this internal ranking of the Google Play Store – in our case for German users – as the metric and acquired the apps on the 27th of June 2019. This list of applications include frequently used streaming service apps, i.e., Spotify, Netflix and Amazon Music, messenger apps, i.e., Snapchat, WhatsApp and LIME, as well as dating applications, i.e., Tinder. We listed an excerpt in Table 1).

The software Raccoon3 (Onyxbits, 2019) is a desktop client for the Google Play Store, which downloads the installable Android application and stores it on the host system. After downloading all applications from the top 100 list, we installed all apps on the Android device via ADB for the subsequent analysis.

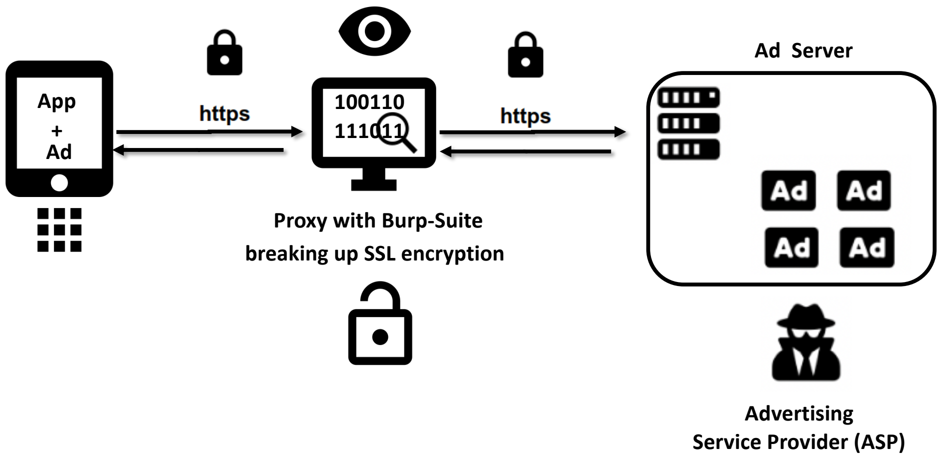
Table 1 This table shows a selection of the German top 100 apps as of 27 June 2019

Joyn	TikTok	TIER	Amazon Shopping
eBay Kleinanzeigen	Spotify	Youtube Music	Wish
Snapchat	Messenger	Instagram	Netflix
MP3 Music Downloader	Enlight Pixaloop	LIME	Deezer
Vova	Circ	WhatsApp	Pinterest
Amazon Music	Tinder	Shazam	AppLock

Note: They include frequently used streaming services as well as messenger and dating apps.

Usually, all connections to the ad servers should use the HTTPS protocol. This secure connection prevents sniffing the network traffic. Thus, we have to setup a proxy as a man-in-the-middle approach on the smartphone and redirect the traffic accordingly. This includes the installation of a self-generated (root) certificate and configuring the device to always use the transparent proxy (Chothia et al., 2017).

Figure 2 The basic setup to analyse the network traffic of Android applications



Notes: We use Burp-Suite to create a man-in-the-middle setup, which enables us to analyse the unencrypted data.

Figure 2 shows an overview of the basic setup: the application tries to establish a secure connection with the ad server by using the HTTPS protocol. We intercept this connection by using a proxy and then relay the data using another secure connection we control ourselves. We used the software Burp-Suite (Portswigger, 2020) as transparent proxy and analysed the network traffic between the app and the advertising network.

Aside from merely inspecting the network traffic, we needed to analyse the binary form of the apps, i.e., to identify the advertisement frameworks used and to analyse the methods they use to acquire privacy-related information. We used a multi-step process, starting with AppBrain Ad Detector (AppBrain, 2019). The application identifies the advertising networks contacted by the original app. With the list of networks, we identified the responsible software libraries in the Android Package (APK) of the app. Finally, we examined each individual app and its internal structure including the decompiled code. For this step, we used the Dex to Java decompiler JADX (JADX-Developer, 2020).

Using this setup, we collected information on the data transferred to from the device to the ad servers and analysed how this data was acquired by the different frameworks. The analysis focuses heavily on the privacy-related information about the user obtained by the frameworks and then passed on to the ASP. We performed three steps in this process: the static code analysis, transmission analysis and checking the data protection declarations of the ASPs against the results.

Please note that the General Data Protection Regulation (GDPR) (European Commission, 2016) is the EU Regulation "on the protection of natural persons with regard to the processing of personal data and on the free movement of such data." It harmonises the rules governing the use of personal data across the European countries. Following the regulation, Article 5 (principles relating to processing of personal data), Article 6 (lawfulness of processing) and Article 7 (conditions for consent) are important for our analysis, as the explicit and legitimate purpose needs to be specified, the minimal amount of data necessary should be collected and – typically – the user's consent is required.

4 Results

Analysis of the top 100 free apps from the Google Play Store revealed that some advertising networks were included in many applications. Others, on the other hand, were contained in very few applications. With this result, the leading advertising networks among the analysed apps could be identified. An overview of the most frequent advertising networks can be found in Figure 3. The leader is the AdMob advertising network provided by Google, which was found in 43 of the 100 applications analysed. The adjust service came in second, closely followed by the Facebook Audience Network, the advertising network of the social media company Facebook in third place. The advertising networks MoPub, Appsflyer and Amazon Mobile Ads were also found in at least ten of the analysed applications.

Figure 3 Advertising networks in the top 100 apps

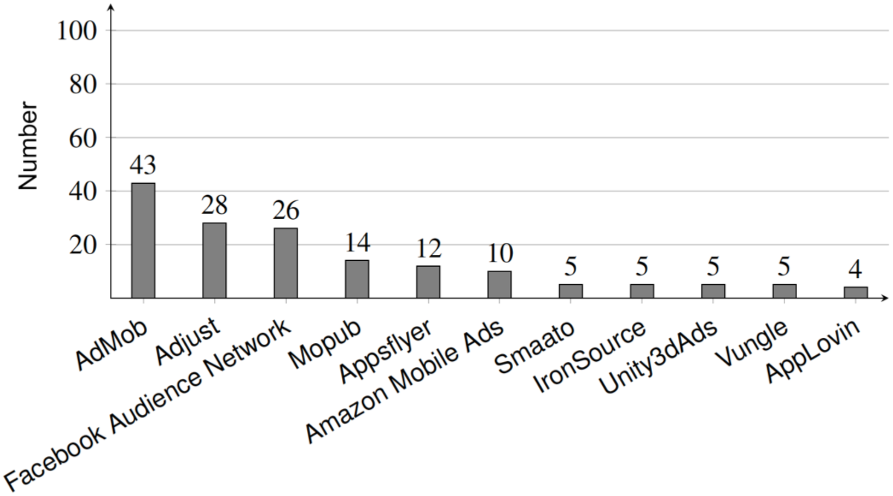


Table 2 Detected device and identifier information

Information		*Adjust*	*AppLovin*	*Appsflyer*	*Facebook Audience Network*	*Ironsource*	*Mopub*	*Smaato*	*Unity Ads*	*Vungle*
Identifier	Android Advertising ID	✓	✓	✓	✓	✓	✓	✓	✓	✗
	Android Device ID	-	-	✓	-	-	-	-	-	-
	MAC-Address	-	-	-	-	-	-	-	?	-
	IMEI	-	-	-	-	-	-	-	?	-
	IMSI	-	-	-	-	-	-	?	-	-
Device	Manufacturer	✗	✓	✗	✗	✓	✓	✗	✓	✗
	Model	✗	✓	✓	✗	✗	✓	✗	✓	✗
	Device type	✗	-	✓	-	-	?	✓	-	✗
	Build number	✗	-	✗	-	-	-	-	-	✗
	Username	-	✗	✗	-	-	✗	-	-	✗
	Operation system	✗	✓	?	✓	✓	✓	✓	✓	✗
	OS version	✗	✗	?	✓	✓	✗	✗	✓	✗
	API-level	✗	✗	✗	-	✗	-	-	✗	✗
	Screen height	✗	-	✗	✗	✗	✗	✗	✗	✗
	Screen width	✗	-	✗	✗	✗	✗	✗	✗	✗
	Screen size	✗	?	✗	-	✗	?	-	✗	-
	Screen resolution	✗	-	-	-	-	-	-	✓	-
	Pixel tightness	-	✗	✗	-	-	-	-	✗	-
	Screen orientation	✗	✓	-	-	✗	-	✗	✗	-
	Aspect ratio	✗		-	-	-	-	-	-	-
	Sound level	-	✓	-	-	✗	-	-	✗	✗
	Battery state	-	?	✗	✓	?	-	-	✗	✗
	Battery load	-		-	✗	✗	-	-	✗	✗
	Device memory (available)	-	?	?	✓	✓	-	-	✗	✗
	RAM total	-	-	-	✗	-	-	-	-	-
	RAM available		-	-	X	-	-	-	-	-
	CPU architecture	✗	-	✓	-	-	-	-	?	-
	CPU number	-	-	-	-	-	-	-	?	-

Table 2 Detected device and identifier information (continued)

Information		*Adjust*	*AppLovin*	*Appsflyer*	*Facebook Audience Network*	*Ironsource*	*Mopub*	*Smaato*	*Unity Ads*	*Vungle*
Device	Available device sensors	-	-	✕	-	-	-	-	-	-
	Motion parameter	-	-	✓	✕	-	-	-	-	-
	Installed apps	-	-	-	-	-	?	?	?	-
	Network connection type	✕	?	✓	✕	✓	?	✕	✕	✕
	Mobile phone operator	-	✓	✓	✓	?	?	?	✕	✕
	User agent	✕	✕	✓	✕	?	✕	✓	✓	✕
	Language	✕	?	✓	?	✓	-	-	?	✕

Table 3 Detected artefacts about location, application and others

Information		*Adjust*	*AppLovin*	*Appsflyer*	*Facebook Audience Network*	*Ironsource*	*Mopub*	*Smaato*	*Unity Ads*	*Vungle*
Location	IP-address	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Country	×	✓	×	-	×	×	-	-	-
	Time zone	-	?	?	?	?	✓	-	×	×
	Locale (i.e., de_DE)	-	×	-	×	-	-	×	×	×
	City	-	-	-	-	-	×	-	-	-
	GPS-coordinates	-	-	-	?	-	?	?	-	-
Application	Name	-	✓	?	✓	?	✓	-	?	-
	Packet	×	×	×	×	✓	×	×	×	×
	Version	×	✓	×	×	×	×	-	×	×
	Installation date	?	×	×	-	-	-	-	-	-
	Time of first execution	?	-	×	-	-	-	-	-	-
	Utilisation time	×	-	?	-	-	?	-	-	-
	Execution within an emulator?	-	-	-	×	-	-	-	-	×
	Size of APK	-	-	-	×	-	-	-	-	-
	Version of advertisement-SDK	-	×	-	×	✓	-	-	-	-
	Type of advertisement	-	-	?	×	-	?	-	-	-
Miscellaneous	Unknown source enabled?	-	-	-	-	-	-	-	-	×
	Personalised advertisements activated?	×	-	×	-	×	×	×	×	×
	Google Play services available?	-	-	-	-	-	-	-	-	×
	Energy saving-mode on?	-	-	-	-	-	-	-	-	×
	SIM-card available?	-	×	-	-	-	-	-	-	-
	Headset plugged in?	-	-	-	-	-	-	-	×	-
	Sound enabled?	-	-	-	-	-	-	-	-	×
	SD card available?	-	-	-	-	-	-	-	-	×
	Root available?	-	-	-	×	×	-	-	-	-
	Data connection throttled ?	-	-	-	-	-	-	-	×	×
	Time stamp	×	-	✓	×	✓	-	?	-	×

Furthermore, the analysis of the apps revealed that the vast majority contained several advertising networks at the same time. The top performer was the app Drum Pad Machine, in which 14 advertising networks could be found. A selection of advertising networks was made on the basis of these results. For this a total of ten advertising networks were selected and further investigated within the scope of the study.

About the results of the analysis, the data sent, and the data provided by the advertising networks, a summary table was drawn up (see Tables 2 and 3). The data transmitted and provided by the advertising networks are marked with a hook ✓. If the data was given, but the transmission could not be determined, these are marked with a question mark (?). Data that were finally transmitted without a specification of the advertising networks are marked with a cross (✕).

4.1 Unique identifiers

During the analysis of all captured network frames, it was found that the character string was included (see Table 2). This is the so-called Android Advertising ID, which is provided for the placement of personalised advertisements through the Google Play Services. This ID is a unique identifier with which data can be assigned to a device and always consists of 32 characters. However, it is possible to reset this ID under the settings of the device.

With this in mind, a search was carried out for other unique identifiers which have a similar benefit. It was determined that the advertising network Appsflyer additionally transmits the so-called Android Device ID. This identifier is created during the installation of the Android operating system and only changed by a new installation. So, even if somebody reset the Android Advertising ID, it could still be recognised by an advertising network with this token. Furthermore, information could be observed by the recorded data, which permits a localisation. This information includes the IP address, which is sent to the server of the advertising network. Furthermore, information about the country and the current timezone was discovered. In the case of the MoPub advertising network, it could also be observed that information about the nearest city was included in the data packet. This information enables an advertising network to place advertisements with a local reference. This is problematic from a privacy perspective because geolocations are specially protected on some legislations.

4.2 More artefacts

Most of the transmitted artefacts were about the device (Table 3). This includes device-specific information such as the model used (e.g., MiA2), the manufacturer (e.g., Xiaomi and Samsung), the operating system name as well as system language and screen dimensions. Information about CPU architecture and the number of CPUs could also be found. Beyond this, the recording also contained status information. These included, among others, current volume settings, battery status (charging, discharging), battery status, available device memory as well as available device sensors, including the motion parameters. The capture of the Vungle advertising network also contained data that indicates the configuration of the Android system, like the available hardware and properties of the mobile connection. Furthermore, information was transferred about the app that was running on the smartphone. This includes the name or process name of the app and the version. In addition, data about the time of installation of a particular app could be discovered for the advertising networks AppLovin and Appsflyer. For Appsflyer the time of first execution could also be found in the capture.

5 Conclusions

In 2018, the French Data Protection Commission (CNIL, 2018) caused quite a stir with a decision in the advertising industry. In their view, data protection supervision not only requires consistent handling when obtaining and passing on consent to third parties, but also considerably more transparency for the user. Accordingly, the French advertising network Vectaury had to revise its system for obtaining the consent of users and passing it on to advertisers. This decision has far-reaching consequences for all advertising network operators within the Europe. The CNIL criticises the fact that users cannot use apps without the SDK being disabled. The inseparable connection between the app and the SDK automatically leads to the transfer of data to the ad server. If companies want to place targeted advertisements via the advertising network, they must inform users which other parties will receive their data. According to the CNIL, it is not enough to inform the user of the terms and conditions. Informed consent requires prior information during installation process.

On 22 May 2019, the Irish Data Protection Commission (2019) opens statutory inquiry against Google Limited. The aim of the statutory inquiry is to determine, whether the processing of personal data carried out at any stage of an advertising operation complies with the relevant provisions of the GDPR.

These two examples show how important the issue is taken by the European authorities. The authors believe that this is a global problem. While the above examples address the legal issues of this problem, the present article focuses on current practice in the top advertising networks.

The results of the study discussed in this article save as an up-to-date overview of the data collection behaviour in advertising networks. More precisely, it could be shown which processes are behind the fade-in of an advertisement. The results show that the on-going discussion on data protection within apps is still a crucial issue. It appears that there is a need for improvement on the part of advertising networks, in particular with regard to the provision of transmitted data. It has been shown that advertising networks transmit more data than they indicate within their privacy policy. The data collected should be specified more precisely in the data protection statements in order to create more transparency and to be really GDPR compliant.

With respect to this study, it was found that identifiers were transferred along with data about the device, approximate location, and information about the application being performed. This again allows a fingerprinting of the smartphone and, therefore, the user. This information cannot directly be used to derive a real identity, but a device and its transmitted data can be uniquely identified across multiple applications. Furthermore, it can be stated that almost all of these data are stored without or via standard authorisations. Both do not require approval.

However, in addition to the presented results, it was observed that even today, program code is still loaded from the internet via the requests of the software libraries. This is still a security risk. The results of this contribution can also be used for future work, for example, to develop a system for classifying advertising networks in Android apps. For this purpose, it would be useful to develop a command-line tool that automatically classifies the contained advertising networks into a more significant number of apps. Using the AppBrain Ad Detector application has been quite complicated for this process, as each app is viewed individually.

Acknowledgements

The project in part has received funding from the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No. 832800.

References

AppBrain (2019) *App Brain Ad Detector Homepage* [online] https://www.appbrain.com/app/appbrain-ad-detector/com.appspot.swisscodemonkeys.detector (accessed 20 November 2019).

Backes, M., Bugiel, S. and Derr, E. (2016) 'Reliable third-party library detection in Android and its security applications', in Weippl, E.R., Katzenbeisser, S., Kruegel, C., Myers, A.C. and Halevi, S. (Eds.): *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, ACM, Vienna, Austria, 24–28 October, pp.356–367.

Book, T., Pridgen, A. and Wallach, D.S. (2013) 'Longitudinal analysis of Android ad library permissions', *CoRR*, abs/1303.0857.

Chothia, T., Garcia, F.D., Heppel, C. and Stone, C.M. (2017) 'Why banker Bob (still) can't get TLS right: a security analysis of TLS in leading UK banking apps', in Kiayias, A. (Ed.): *Financial Cryptography and Data Security – 21st International Conference, FC 2017, Lecture Notes in Computer Science*, Revised Selected Papers, Springer, Sliema, Malta, 3–7 April, Vol. 10322, pp.579–597.

Commission Nationale de l'Informatique et des Libertés (CNIL) (2018) 'Décision no MED 2018-042 du 30 Octobre 2018 mettant en demeure la société VECTAURY', *Détail d'une Délibération de la CNIL*.

Crussell, J., Stevens, R. and Chen, H. (2014) 'MAdFraud: investigating ad fraud in Android applications', in Campbell, A.T., Kotz, D., Cox, L.P. and Mao, Z.M. (Eds.): *The 12th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys'14*, ACM, Bretton Woods, NH, USA, 16–19 June, pp.123–134.

Data Protection Commission (2019) *Data Protection Commission Opens Statutory Inquiry into Google Ireland Limited* [online] https://www.dataprotection.ie/en/news-media/press-releases/data-protection-commission-opens-statutory-inquiry-google-ireland-limited (accessed 13 June 2019).

Dogruel, L. (2019) 'Too much information!? Examining the impact of different levels of transparency on consumers' evaluations of targeted advertising', *Communication Research Reports*, Vol. 36, No. 5, pp.383–392.

Esteve, A. (2017) 'The business of personal data: Google, Facebook, and privacy issues in the EU and the USA', *International Data Privacy Law*, Vol. 7, No. 1, pp.34–47.

European Commission (2016) *General Data Protection Regulation (GDPR) 2016/679* [online] https://eur-lex.europa.eu/eli/reg/2016/679/oj (accessed 13 February 2020).

Federal Trade Commission (2016) *Mobile Advertising Network InMobi Settles FTC Charges It Tracked Hundreds of Millions of Consumers' Locations Without Permission* [online] https://www.ftc.gov/news-events/press-releases/2016/06/mobile-advertising-network-inmobi-settles-ftc-charges-it-tracked (accessed 16 November 2020).

Federal Trade Commission (2020) *Making Sure Companies Keep their Privacy Promises to Consumers* [online] http://www.ftc.gov/news-events/media-resources/protecting-consumer-privacy/enforcing-privacy-promises (accessed 3 December 2020).

Google Developers (2020) *Android Debug Bridge (ADB) – Android Studio User Guide* [online] https://developer.android.com/studio/command-line/adb (accessed 22 December 2020).

Google Inc. (2020) *Google Privacy Policy* [online] https://support.google.com/adspolicy/answer/54817?hl=en (accessed 1 December 2020).

Grace, M.C., Zhou, W., Jiang, X. and Sadeghi, A. (2012) 'Unsafe exposure analysis of mobile in-app advertisements', in Krunz, M., Lazos, L., Pietro, R.D. and Trappe, W. (Eds.): *Proceedings of the Fifth ACM Conference on Security and Privacy in Wireless and Mobile Networks, WISEC 2012*, ACM, Tucson, AZ, USA, 16–18 April, pp.101–112.

Graef, I. and Berlo, S.V. (2020) 'Towards smarter regulation in the areas of competition, data protection and consumer law: why greater power should come with greater responsibility', *European Journal of Risk Regulation*, pp.1–25.

JADX-Developer (2020) *JADX – Dex to Java Decompiler*, Offical GitHub Page [online] https://github.com/skylot/jadx (accessed 21 February 2020).

Kees, J. and Andrews, J.C. (2019) 'Research issues and needs at the intersection of advertising and public policy, *Journal of Advertising*, Vol. 48, No. 1, pp.126–135.

Kim, T., Barasz, K. and John, L.K. (2018) 'Why am I seeing this ad? The effect of ad transparency on ad effectiveness', *Journal of Consumer Research*, Vol. 45, No. 5, pp.906–932.

Lian, S., Cha, T. and Xu, Y. (2019) 'Enhancing geotargeting with temporal targeting, behavioral targeting and promotion for comprehensive contextual targeting', *Decision Support Systems*, Vol. 117, pp.28–37.

Narayanan, A., Chen, L. and Chan, C.K. (2014) 'AdDetect: automated detection of Android ad libraries using semantic analysis', *2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, IEEE, Singapore, 21–24 April, pp.1–6.

Onyxbits (2019) *Raccoon3 the APK Downloader* [online] https://raccoon.onyxbits.de/ (accessed 20 December 2019).

Portswigger (2020) *The Burp Suite Family* [online] https://portswigger.net/burp (accessed 22 December 2020).

Son, S., Kim, D. and Shmatikov, V. (2016) 'What mobile ads know about mobile users', *23rd Annual Network and Distributed System Security Symposium, NDSS 2016*, The Internet Society, San Diego, California, USA, 21–24 February [online] http://wp.internetsociety.org/ndss/wp-content/uploads/sites/25/2017/09/what-mobile-ads-know-about-mobile-users.pdf.

Stevens, R., Gibler, C., Crussell, J., Erickson, J.L. and Chen, H. (2012) *Investigating User Privacy in Android Ad Libraries*, University of California.

Vallina-Rodriguez, N., Sundaresan, S., Razaghpanah, A., Nithyanand, R., Allman, M., Kreibich, C. and Gill, P. (2016) 'Tracking the trackers: towards understanding the mobile advertising and tracking ecosystem', *CoRR*, abs/1609.07190.

Yuan, S., Abidin, A.Z., Sloan, M. and Wang, J. (2012) 'Internet advertising: an interplay among advertisers, online publishers, ad exchanges and web users', *CoRR*, abs/1206.1754.