Association for Information Systems
AIS Electronic Library (AISeL)

MCIS 2017 Proceedings

Mediterranean Conference on Information Systems (MCIS)

9-2017

Mobile Application Privacy Risks : Viber Users' De-Anonymization Using Public Data

Ioannis Paspatis
Ionian University, c16pasp@ionio.gr

Aggeliki Tsohou
Ionian University, atsohou@ionio.gr

Spyros Kokolakis
University of the Aegean, sak@aegean.gr

MOBILE APPLICATION PRIVACY RISKS: VIBER USERS' DE-ANONYMIZATION USING PUBLIC DATA

Research full-length paper

Track 10- Trust, Security and Privacy Track

Paspatis, Ioannis, Ionian University, Corfu, GR, c16pasp@ionio.gr

Tsohou, Aggeliki, Ionian University, Corfu, GR, atsohou@ionio.gr

Kokolakis, Spyros, University of Aegean, Samos, GR, sak@aegean.gr

Abstract

Mobile application developers define the terms of use for the applications they develop, which users may accept or declined during installation. Application developers on the one hand seek to gain access to as many user information as possible, while users on the other hand seem to lack awareness and comprehension of privacy policies. This allows application developers to store an enormous number of personal data, sometimes even irrelevant to the application's function. It's also common that users choose not to alter the default settings, even when such an option is provided. In combination, the above conditions jeopardize users' rights to privacy. In this research, we examined the Viber application to demonstrate how effortless it is to discover the identity of unknown Viber users. We chose a pseudorandom sample of 2000 cellular telephone numbers and examined if we could reveal their personal information. We designed an empirical study that compares the reported behavior with the actual behavior of Viber's users. The results of this study show that users' anonymity and privacy is easily deprived and information is exposed to a knowledgeable seeker. We provide guidelines addressed to both mobile application users and developers to increase privacy awareness and prevent privacy violations.

Key words: Viber, privacy, awareness, anonymity, privacy deprivation, public data, mobile apps.

1 Introduction

The massive increase of smartphone use has created new opportunities in the telecommunications' field. Multiple instant messaging applications like Viber, WhatsApp and Facebook Messenger are being downloaded, installed, and used daily to enhance human communication, making these apps an important part of almost anyone's social life. According to the surveys by Statista (Statista, 2017a; 2017b; 2017c) in 2017's first quarter, there were 1.2 billion WhatsApp and Facebook Messenger and 891 million Viber's unique users. Such mobile applications (hereafter mobile apps) have evolved from being a simple instant messaging service to a multi-use social networking tool. Apart from Instant Messaging, they provide voice messaging, file and photos sharing, location tracking, voice calls and many more features. Mobile app developers (hereafter app developers) are responsible for the mobile app's permission set. This allows app developers to collect a vast amount of information on users, such as mobile numbers, locations, contact lists, profile pictures, or other users' private information and to use them at will or opportunity, including selling data to third parties.

Recent studies have concluded that users consent to the mobile app's permission set, as a mandatory step in the downloading process, without actually reading the privacy policy. The result is the misinformed exposure of their personal data. McDonald and Cranor (2009) illustrated that such privacy policies are far too long and complicated for the users to spend their time reading them. Giovani and Pashley (2007) state that even though users are informed about the possible consequences of a leakage of personal data and the imminent dangers involved, they don't change their security settings in their social networks and they keep on using them as usual, although they know how to change them. According to Tow et al. (2008), users are not well informed and, therefore, not aware of their personal data exposure, or they feel it is unlikely that such a privacy incident could harm them. So, users are unaware of the risks they are facing, such as identity forge, stealing, impersonation and many others, as they lack the knowledge required to evaluate these risks (Pitkanen & Tuunainen, 2012).

Our aim is to demonstrate how effortless it is to discover the identity of an unknown app user by only processing self-exposed public data and to offer guidelines for mobile app users and developers. We revealed users' personal information such as their real name, address and other personal data. It is crucial to show that even without any prior information we can extract automatically information on one's private habits, such as the time they wake-up or when they switch on and off their mobile phone. More importantly, we can do this for a large number of people, and not only for targeted individuals. The purpose of such a demonstration is to enhance privacy awareness of mobile app users and to inform them about the risks involved in disclosing their public data, so that they would change the privacy settings on their mobile apps and stop blindly accepting privacy policies.

In order to examine our claim that users' identity can be revealed by publicly available data, we conducted an empirical experiment using the popular instant messaging mobile app Viber. We chose a set of 2000 pseudo-random cellphone dummy numbers and we automatically discovered which of them have the Viber app installed. We had a prior knowledge that these cellphone numbers were active in the recent past. The de-anonymization process was successful for 75% of our test sample. In most cases, we disclosed the full name of the subject and his/her address. This study is the first, to the best of the authors' knowledge, to show that this kind of information can be gathered in a large scale without targeting specific subjects. Buchenscheit et al. (2014) have similarly studied WhatsApp mobile app and concluded through an experiment that it is possible to infer users' habits. Although the study of Buchenscheit et al. (2014) reveals several privacy related inferences, it does not show that users' identities can also be revealed, which we do in this paper.

The paper is structured as follows. In section two we present the related work. Section three presents the methodology we followed and the preparation we made in order to accomplish our data mining aims and create meaningful statistic data. In section four, we present our de-anonymization results, our statistical analysis and our empirical questionnaire results (see Appendix). Section five presents the contribution of the paper and a guide for best practices to reduce private data leak. Finally, section six concludes the paper.

2 Related Work

In order to examine scholar disquisition associated to our research objective, we reviewed the following literature areas: (a) general literature on personal data disclosure, (b) users' privacy awareness, and (c) mobile apps' features and vulnerabilities. At section 2.3, we specifically examined the Androids OS security characteristics and features. We chose to study Android OS because it allows users to change the security settings (Benenson & Reinfelder, 2013). On the other hand, IOS uses an App review process system when a new app is uploaded to its App Store, whilst users cannot change any security settings or review the permission set before installation.

2.1 Personal Data Disclosure

The invasion and violation of one's privacy and personality includes unauthorized collection, publishing and/or posting of personal data through the Internet and any other use of personal information without the prior consent of the data subject (Wang et al., 1998). The collection, sale and/or exchange of personal data between companies or individuals has become a common practice (Gillmor, 1998) as the Internet has made it easy to collect and store large volumes of personal information (Caruso, 1998). According to a 2002 survey (Rust et al., 2002) it has become impossible for consumers to carry out almost any electronic transaction without revealing some sort of personal information. Several online companies are operating on collecting and using such personal data (Fletcher, 2003). Users show an increased concern when their data are collected and disclosed on the web (Global Internet User Survey, 2012; TRUSTe, 2014). Several researchers have set out indicators on consumers' concerns when the handling of their personal information is involved (Sheehan and Hoy, 2000; Dinev and Hart, 2004; Bellman et al., 2004). Users' awareness on data collection, misuse of data, experience in internet use, social profile and consumers' level of education are some of these indicators.

2.2 Users' privacy awareness and the cost of reading privacy policies

McDonald and Cranor (2008) examined the reasons why users' do not read the privacy policies, as well as time and financial costs of reading them. They investigated 75 of the most popular websites, and they reveal that a user would need about 10 minutes to read one privacy policy, at an average read rate of 250 words/minute. Also, they concluded that an average user should read at least 1354 privacy policies per year, where 412 of them should be read while at work hours (McDonald and Cranor, 2008). Based on the above, they conclude that a user would have to spend 40 minutes per day for just reading policies, while the average user is browsing the internet for approximately 72 minutes per day. Translating this into financial cost would mean about 3500 dollars per year for a user, while at a national level (in the U.S., where the research was conducted) this would be 781 billion per year. Capistrano and Chena (2015) showed that the specificity of the privacy policy has a significant effect on consumers' perceptions of its importance in deciding to share personal information. The key factor for their decision was the length of the privacy policy as primary reason as well as the visibility of its.

Govani and Pashley (2007) focused on the privacy awareness of university students in relation to the Facebook social network. According to their findings the majority of students are aware of the potential impact from the leakage of personal data as well as the imminent danger of impersonation or monitoring. Although they are aware of their options to limit the leakage of personal data, they do not take any action. This is in line with the findings of Tow et al. (2008) that users are either not fully aware of the privacy incidents or feel that the likelihood of privacy violation is too small to happen. Similar results have been shown by Gross and Acquisti (2005). Their findings stated that the risk of personal exposure or identity theft and other negative effects is proportional to the amount of information a user exposes. Almuhimedi et al. (2014) concluded that most users are unaware that their personal data are collected by mobile apps and also that their location can be exposed by these data. In addition, they suggested the use of an app permission manager such as AppOps, an embedded app manager in Android 4.2. The permission manager will notify users with privacy nudges in case an app starts to use a permission that

can expose their personal information. Similar privacy nudges are present iOS 6 and after, where users receive popup warnings about usage of some data types and are asked for informed consent (Benenson and Reinfelder, 2013). As conclusion of the above, is proved that the users' low awareness can lead to privacy risk incidents or increase the chance for them to be happen. Researchers also concluded that users avoid reading privacy policies due to their extended length or the time need it for it.

2.3 Mobile apps' features and vulnerabilities

The Mobile Device Safety Management with Android operating system is implemented with a multisite security system (Boksasp et al., 2012). This Safety Management System includes: (a) app's permission management system (permission set), (b) sandboxing, (c) unique certification signing for every developer, (d) remote kill switch for detected malfunctioning apps, (e) core file system protection in read only mode, (f) google bouncer as anti-virus, and (g) third-party antiviruses. These security characteristics tend to protect only the device and not user's data. According to a Pew Research Center survey (2012), 82.56% of the app's that are available for downloading in app markets are requiring the permission "Full network access", while 23.75% of apps are requiring the permission "Precise Location GPS and network based Permission". These permissions are giving the opportunity to an app developer to disclose users' real location. Combined with the fact that most users install at least 10 apps on their mobile device (Olmstead & Atkinson, 2015), it creates favorable conditions for revealing users' location. Rakib and Ho (2011) concluded that service providers should consider users' privacy concerns before they use capabilities such as location detection or personal-data detection. Buchenscheit et al. (2014) have studied WhatsApp mobile app and found, through an experiment, that it is possible to infer users' habits. Several researchers have shown that many users not pay attention to the permissions are granting to an app upon installation (Felt et al., 2012; Kelley et al., 2012). The researchers also, concluded that the technical language in which they are written is not easy to understand. Benenson and Reinfelder (2012) showed that Android users would be better informed about possible risks in case they notice and can understand the permissions. The same researchers concluded that although the Android's permissions visibility and cryptic language is difficult to understand, they seem to work in raising awareness at least with more technically savvy users. Appelman et al. (2011) conducted a research regarding Viber's communication security. Some of their findings were: Viber's database is unencrypted and easy accessible, it is possible to bypass the official registration process and one could forge a person's Viber account, data messages are probably sending scrambled and not encrypted.

Users' low privacy awareness and the difficulties they face on reading privacy policies, as mentioned in chapters 2.1 and 2.2, combined with the mobile devices' security management system, create appropriate conditions for the non-consented disclosure of personal information. In our experiment, we show that it is possible to disclose users' personal information, such as real name and surname, profile photos and address without any prior knowledge of their identities. This study is the first, to the best of the authors' knowledge, where these vulnerabilities of the security management system are explored with actual personal information for empirical investigation.

3 Preparation and methodology

Viber's users are nearly 900 million as of March 2017. This makes Viber a perfect candidate to conduct our research. There are three core characteristics of this app that are essential to note. Firstly, when a user installs Viber, her cell phone number is sent to Viber's servers by default. Also, as feature, the contact list of the new user's device is parsed, and every Viber's user who is registered at this contact list is automatically informed about the new user enrolment. Secondly, there is a practical difficulty for a user to exit the mobile app. Even if the user presses the exit button from the app's menu, the app will continue running in the background as a service. Lastly, the mobile app allows a massive import of contact data in Viber's contact list without any confirmation from the other users. This feature was exploited in our research.

3.1 System Preparation

To create a controlled setting for our experiment, we created a virtual machine with Microsoft Windows 10 and we installed our toolset. Our toolset is composed by Viber for Desktop, DB Browser for SQLite for the mobiles' database handling, Python for our scripts and Netbeans for Java to create our crawlers. The virtual machine is equipped with an Intel i7 6700, 16 GB of RAM and 250 GB SSD SATA 3 disk. Also, a sim card was purchased for the testing device to activate the "Viber for Desktop" application.

3.2 Methodology of data mining

To start our data mining we first wrote an 8-line python script with which we created a 2000-line dummy csv file. The file was grafted as follows: name as dummyName[serial number], surname as dummySurname[serial number], cellphone number as +3069xxxxx[serial number]. Then, we created a new Gmail account and we imported our dummy csv file as real contacts. Next, we installed Viber for cellphones and we added the Gmail account we created. The cellphone numbers were pseudorandom numbers. We had prior knowledge that these phone numbers were employees' phone numbers of a telecommunication company that stopped functioning since 2013. In sequence, the Viber app started to synchronize our dummy dataset with real Viber accounts. After this step, we can see in our testing device which contacts have an account on Viber. This was performed by opening the viber.db with the SQLite app and running the SQL query "select * from Contact where ViberContact = 1 and order by Number". The result includes the contacts who have a Viber account (ViberContact = 1), a possible name and a possible profile picture (DownloadID not null) (Figure 1). These data were exported to a csv file for later processing.

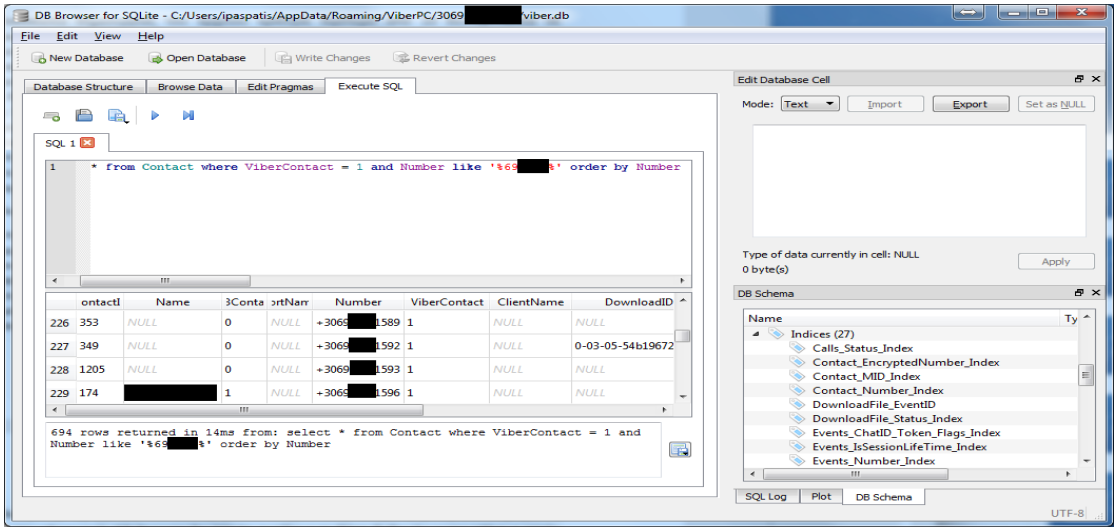

Figure 1: Screenshot from SQLite during datamining

3.3 Methodology of the de-anonymization of the test sample

With the intention to de-anonymize our data set we used two methods. In the first method, we used three public and free databases, Greekphones, Sync.me and TrueCaller. These databases contain freely given personal data from various sources, such as public data records or personal data that were disclosed on social networks such as Facebook, LinkedIn and Twitter (Loie Favre, 2014). We should notice that Sync.me is also using another method to collect data. When Sync.me is installed in a cellphone it uploads all phone numbers that are stored in the device's contact list or in the Gmail account contact list (Heather Clancy, 2013). For the database "Greekphones" we created a custom-made crawler based on java and the jsoup library. Our crawler automatically parses the Name and Surname, the address, the postal code

and, in some cases, the occupation of a given set of Greek phone numbers. The other two databases use the captcha authentication system and limit the time of use. For this reason, we conducted a semi-automatic parse: (a) we connected to these two databases through the Tor network, in order to override the limited times of use, (b) we authenticated manually at least one time, and (c) we used a custom-made java program to find if a given number that wasn't identified with our first crawler, exists in the other two databases. In case of a query limit lockout, we created a new Tor identity and re-authenticated through captcha. For our second method, we used the google reverse image lookup. We created an ftp server and uploaded the profile pictures we discovered during our data mining exploration. After, we created a java program for data handling and used Google reverse image API as an imported library. The program tried to identify a person using the hyperlink profile image provided by our ftp server. If there were similar pictures, then a web browser would open a webpage with these similar pictures for manual identification. Unfortunately, Google has discontinued officially supporting this API (Google, 2011) and since November 2016 it prevents the automatic parse of the findings in aim to promote its new Google Custom Search API. This API limits the searches to 100 per day and, then, charges each search based on the number of results (Google, 2016). If a person used the same profile picture in one or more social networks, then we were able to discover the name of that person.

We distinguish two types of de-anonymization. First, full de-anonymization (or de-anonymization) is considered when we disclose at least the name and the surname of an entity simultaneously. Second, semi de-anonymization occurs when we disclose the name or the surname or a nickname for an entity.

Both data mining and the de-anonymization can be viewed in greater detail in figure 2. We separated figure 2 in ten steps as follow:

Step 1: we created a pseudo-random .csv with a simple python script

Step 2: we uploaded the .csv file to a gmail account

Step 3: In this step the gmail contacts synchronized with our lab smartphone

Step 4: we synchronized the smartphone with our lab pc

Step 5: we mined the data from the sqlite DB using DB browser and simple sql scripts

Step 6: we prepared our databases for handling

Step 7a-8a: we uploaded the viber images that we mined from step 5 to an ftp server in order to do our reverse image lookup

Step 7b-8b: we used our java custom parser to get public information (name, surname, address) about our dataset's subjects

Step 9: we reviewed our results from step 8a and manually we removed the inconvenient data.

Step10: we merged our results to a datasheet file and we conducted our statistical analysis.

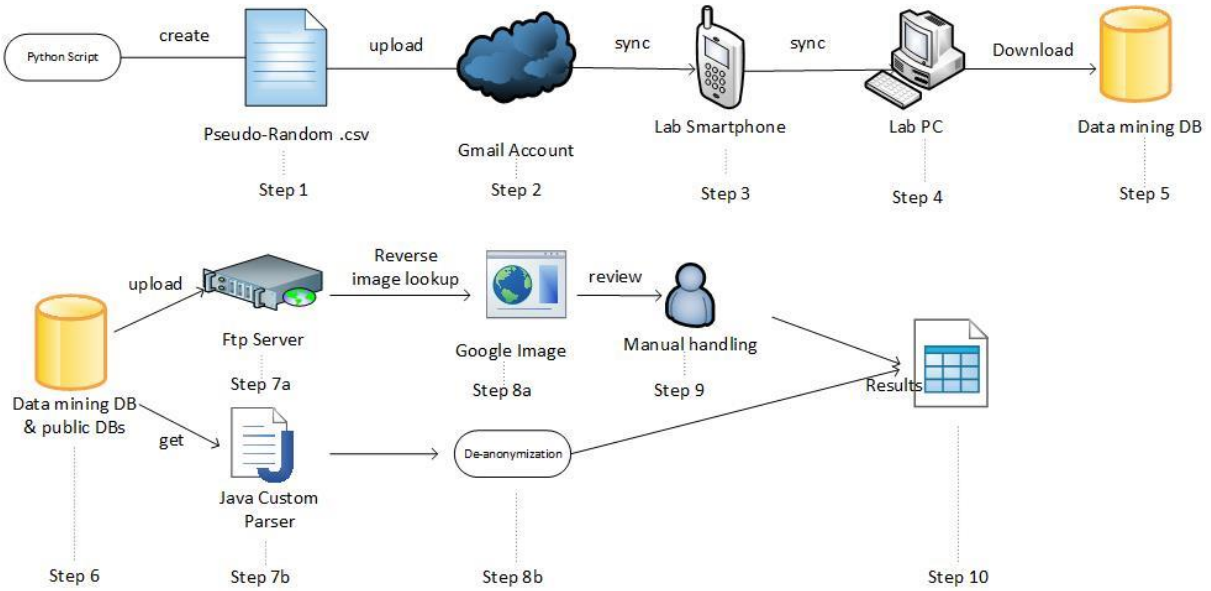

Figure 2: Steps of data de-anonymization

3.4 Methodology of Privacy Awareness Exploration

Within our research we also investigated the privacy awareness level of Viber users regarding the mobile app's privacy settings. For this reason, we designed an empirical study that compares the reported behavior with the actual behavior of Viber users. We followed the approach of convenience sampling and invited participants who were accessible to the researchers. All participants were holders of a computer science diploma. Our sample included 20 individuals, who are users of the Viber mobile app and agreed to participate, being aware of the monitoring processes that we followed.

In order to monitor actual behavior, we included the 20 users in the dataset for 30 consecutive days. The purpose of the survey was to investigate if the Viber's users are aware regarding the features of Viber, such as the possibility of personal data and habits exposure. Also, to explore if they were aware of the potential risks of the blind consent to the privacy policies. The purpose of the monitoring was to determine if we could expose through Viber some of our subjects' habits, such as the time they turn on/off their device or other habits. The monitoring was taking place three times a day at 8.00am, 4.00pm and 10.00pm.

In order to record reported behavior, we distributed a questionnaire that we designed using Google Forms due to its characteristics (e.g., speed of processing the results, security and ease of use). The questionnaire was distributed online via e-mail to our sample. The survey was running for 30 days and all participants replied to the questionnaire.

4 Results and Statistical analysis

4.1 De-anonymization results

After we recovered the accumulated data, we had appropriate and sufficient data to proceed to statistical analysis. From the total of 2000 pseudo-random cellphone numbers it emerged that 682 subscribers have installed Viber for mobile (34%). From the 682 entities, the 316 entities (46.3%) have uploaded a profile picture. From the entities that have uploaded a profile picture we have de-anonymized or semi de-anonymized 258 of them (82%) while from the entities that hadn't upload a profile picture we have de-anonymize or semi de-anonymize 291 of them (80%). In total, from both categories we have de-

anonymize or semi de-anonymize 549 entities. To conclude, we have fully de-anonymized 475 entities (75%). In table 1 we present our results in greater detail.

Description	Sample analysis
Entities with Viber installed (n=2000)	682 (34%)
Entities with profile picture (n=682)	316 (46%)
De-anonymized entities with profile picture (n=316)	258(82%)
Full de-anonymized entities with profile picture (n=258)	220 (82%)
Semi de-anonymized entities with profile picture (n=258)	38 (18%)
De-anonymized entities without profile picture (n=366)	291 (80%)
De-anonymized entities without profile picture (n=291)	255 (88%)
Semi de-anonymized entities without profile picture (n=291)	36 (12%)
De-anonymized entities with or without profile picture (n=549)	475 (87%)
Semi de-anonymized entities with or without profile picture (n=549)	74 (13%)
De-anonymized entities in total (n=682)	549 (80%)
Fail to de-anonymize entities (n=682)	133 (20%)

Table 1. Results of the sample de-anonymization process

4.2 Empirical questionnaire results

Our findings from the online questionnaire reveal that Viber users have a low level of privacy awareness regarding the possibility of personal data exposure and habits' inference, as well as the potential risks of the blind consent to Viber's privacy policy. From the perspective of privacy awareness, our findings showed that all our subjects - except one - never changed their profile picture (95%). 90% of our subjects answered that they didn't read, or didn't remember if they read, the privacy policy before installing Viber. 85% stated that even if they had read it, they weren't sure if they would have understood it. 90% stated that they didn't know that the data they publish on Viber is considered as public data. The same percentage stated that they didn't know that Viber, according to its privacy policy, can share users' public data with other cooperating social networks, like Facebook or Twitter. 50% stated that they find it at least difficult to change Viber's privacy settings. 65% of our subjects stated that they didn't know it's possible to identify someone through reverse image lookup, while 75% answered they didn't know that it was possible for someone to be de-anonymized via her phone number. 60% stated that they didn't know that through Viber they could be a target of spam messages or phishing attacks nor that they could be a target of physical or electronic monitoring. Nevertheless, 35% would suggest to a friend to install Viber, while 60% will continue to use it.

4.3 Daily Monitoring Results

From our daily monitoring of the 20 subjects, the following findings were obtained: The 74.9% of our participants had been connected to Viber at 8.00 am (+/- 2 hours) at all days. The percentage increased during working days at 82.9% (+/- 2 hours). From this, we conclude that most our subjects preferred to connect to Viber from their work space. At 4.00 pm and 10.00 pm the percentage of connected subjects was approximately at 87% the most of the days. From a privacy perspective, we can conclude that our

subjects are more vulnerable to spam messages or unwanted advertising phone calls within this time. Our findings are analyzed in greater detail in the following figures (3, 4). Also, we monitored their profile pictures changes. Only one of the subjects changed his/her profile picture (5%). From the side of privacy this is positive. If the de-anonymization process fails to identify a person through reverse image lookup in first place, it is less probable that (s)he will be identified with this process in the future.

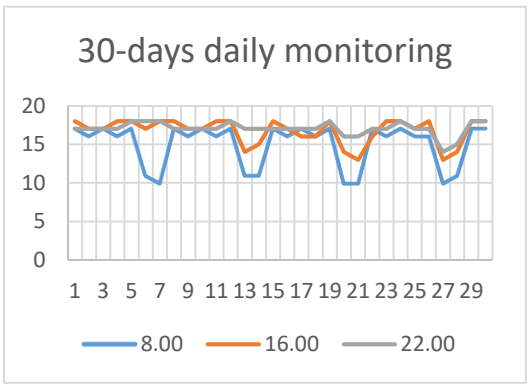

Figure 3: Daily monitoring results

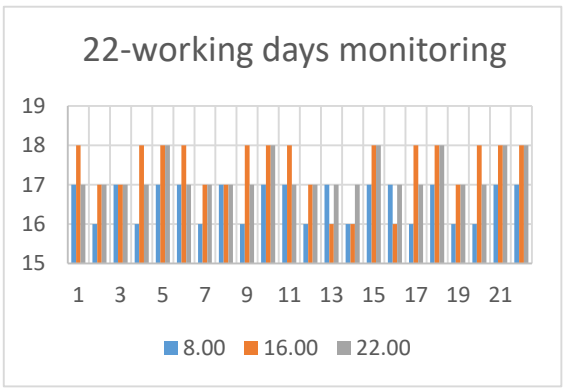

Figure 4: Working days monitoring results

From our findings, we conclude that computer science knowledge is not sufficient to enhance individual privacy awareness. Most our subjects didn't read Viber's privacy policy. Even if they had read it they fear that they wouldn't understand it. The fact that almost all our subjects did not know that the data they share with Viber is considered to be public data demonstrates a basic lack of privacy awareness with regard to personal data. Finally, since they do not change their privacy settings, they can be an easy target for many potential threats, such as spam messaging or monitoring.

5 Contribution and guidelines

5.1 Contribution to the Research community

The importance of users' awareness has been pointed out by the literature review in section 2 and by our experimental and survey results. We proved that even without any prior information about the records of our pseudo-random dataset, we could de-anonymize 75% of our subjects, revealing their names, surnames and/or home addresses. It is important to consider that interested companies (e.g., advertisers or public data handlers) perform such de-anonymization and the derived information is a useful tool for their practices. When it comes to the user's privacy, it appears to be difficult for the average user to keep their personal data from becoming public due to lack of awareness. Therefore, this exposes them to various types of threats, such as physical and electronic monitoring, hacking or spam messaging. The above threats can be abridged by increasing user's privacy awareness, in conjunction with the development of appropriate mobile programming protocols for the developers.

5.2 Guidelines for the Privacy Regulation Authorities regarding mobile app developers

Our research has demonstrated the ease of acquiring personal data using mobile app characteristics. Our experiment would not have been possible if some personal data processing principles had been followed by the mobile app. We suggest that the personal data proception authorities consider our findings and enforce a set of principles for mobile developing companies and mobile app developers, in order to protect the mobile app users:

- Any mobile application should request the users' consent whenever someone tries to add them to their contact list. Even if some mobile apps use implicit consent, we argue that explicit consent is imperative for the provision of this right by a user. Therefore, the default setting for a given user should be to disable the automatic connection with others, and it be at the discretion of users to enable the feature.
- When the mobile app settings allow the user to expose personal data, such as profile pictures or names, the default settings should disable the exposure and it should be at the users' discretion to activate it.
- When the mobile app collects information that may allow the exposure of personal habits (e.g., state the device, turned on or off) the default settings should disable the collection of such data and be activated only after users' explicit consent. The users should be warned about the potential risks they may face.

5.3 Guidelines for mobile app users

From our experiment results it is shown that it is possible to expose users' personal data or habits regarding the features of a mobile app. Our experiment wouldn't have succeed if an informed user had disabled some of these characteristics. Below we provide guidelines that will help users to protect their data.

To mobile app users we recommend:

- Always read the privacy policy before accepting it. It may contain valuable information about your personal data handling.
- Read the mobile app permissions that are granted during the installation. Disable the permissions you consider as dangerous even if this limits the mobile app's functionality (this feature is available only from Android version 7 and on).
- Check the mobile app's privacy settings before you use it or add your personal information. They may contain features that may expose your personal data, such as usual or temporary location, name, etc.
- We recommend to never upload a real profile picture. It can be used to de-anonymize you or to link your entity to the social networks you may use.
- If there is a setting that prevents someone to add you to a social network without your consent, consider to enable it. Since many personal data crawlers use auto-join features to profile you, your authorization to such an act, can be proved useful to keep your information private.

6 Conclusions

In this paper, we have proved that it is possible to de-anonymize Viber users without any prior knowledge about their identity. We selected a pseudorandom sample of phone numbers and successfully de-anonymized 75% of 682 Viber users by acquiring and disclosing their real name, surname and address. Our experiment was successful due to some low security characteristics of Viber, that have been reported also by Appelman et al. (2011). We have applied software tools for the data mining process, using only freeware tools and software we developed ourselves and de-anonymize our dataset with custom made crawlers in conjunction with three, free to use, online phone databases. From the above, we conclude how important is to raise users' awareness for privacy implications in Viber and other mobile apps in order to protect their personal data.

We investigated the privacy awareness level of Viber users regarding the general features of Viber app that could expose their personal data, as well as the importance of reading the privacy policies. Our exploratory survey included 20 Viber users as participants, all of them holding a computer science

diploma. The results verified our initial conjecture about the low level of privacy awareness regarding the general features of Viber that could expose their personal data and put themselves in potential risk. Indeed, most of the participants were not aware that they could be identified with the use of online databases, that their personal habits could be monitored or that their private data is considered as public data after they disclose them, voluntarily on mobile applications. Despite the high level of computer science education of the participants, they weren't tempted to read Viber's privacy policy or alter Viber's privacy settings. Our results emphasize the need to raise privacy awareness of Viber users and of mobile applications, in general.

Based on our findings, we provided general guidelines for the personal data regulation authorities and more specific recommendations for the profession of mobile apps developers, all of which aim to protect the personal data of mobile apps users by default. We also offered specific guidelines to mobile apps users with the main intention to help them protect their personal data, especially since misuse of their data could put them at risk.

7 Appendix

A. Demographics

1. Gender?

   Male\Female

2. What is your age?

   17 or younger\18-20\21-29\30-39\40-49\50-59\60+

3. What is the highest level of school you have completed or the highest degree you have received?

   Less than high school degree\High school degree or equivalent (e.g., GED)\Some college but no degree\Associate degree\Bachelor degree\Masters degree\P.H.D. Graduate degree

4. Do you have an account on a social networking website (like Facebook or Twitter)?

   Yes\No\Maybe

5. About how many contacts do you currently have on social networking websites?

   0-100\100-500\500+\I have no social network profile

6. About how many of your "friends" on social networking websites have you met in person?

   All of them\Most of them\About half of them\A few of them\None of them

7. How often do you change your profile picture on social networking websites?

   Extremely often\Very often\Moderately often\Slightly often\Not at all often

8. How easy do you think it is to modify your settings on your social networking websites?

   Extremely easy\Very easy\Moderately easy\Slightly easy\Not at all easy

9. If you could use only one of the following social networking services, which would you use?

   Facebook\Google+\MySpace\Viber\Twitter\Whats App

10. When you're on social networking websites, about how much of your time do you spend looking at what other people have posted?

    All of it\Most of it\About half of it\Some of it\None of it

11. When you're on social networking websites, about how much of your time do you spend posting things about yourself?

    All of it\Most of it\About half of it\Some of it\None of it

B. What about Viber (Awareness)

1. Had you read the Privacy Policy Before you installed Viber?

Yes\No\Maybe/I can't remember\I don't want to answer

2. If you read the Privacy Policy did you understand it?

Probably No\Yes\No\I'm not sure\I don't want to answer

3. Do you know that by installing viber your viber data are now considered as public data?

Yes\No\I don't know\I don't want to answer

4. Do you know that according to Viber's Privacy Policy Viber can share your data with third parties e.g. social networks like Facebook and Twitter?

Yes\No\I don't know\I don't want to answer

5. Do you know that Viber can collect data due to the agreement with other social networks like Facebook and Twitter?

Yes\No\I don't know\I don't want to answer

6. Do you know that someone could find your profile picture and other personal data like your full name or/and your profession/occupation?

Yes\No\I don't know\I don't want to answer

7. Do you know that strangers, employers or friends can follow your habits through Viber e.g. what time you switched on your mobile device or check your connection status?

Yes\No\I don't know\I don't want to answer

8. Do you know that through Viber you can be a target of spam messages?

Yes\No\I don't know\I don't want to answer

9. Do you know that through Viber you can be a target of electronic phishing?

Yes\No\I don't know\I don't want to answer

10. Do you know that by using linked data from multiple social networks such as Viber, Twitter, Facebook you may become electronic and physical monitoring target?

Yes\No\I don't know\I don't want to answer

11. After the above will you recommend Viber to a friend?

Yes\No\I don't know\I don't want to answer

12. Will you continue using Viber?

Yes\No\I don't know\I don't want to answer

References

Almuhimedi H, Schaub F, Sadeh N, Adjerid I, Acquisti A, Gluck J, et al. (2014). Your location has been shared 5,398 times! A field studyon mobile app privacy nudging. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 787–796.

Appelman M, Bosma J, Veerman G, (2011). Viber Communication Security. Unscramble the scrambled. System and network of engineering, University of Amsterdam. [online]. Available at: https://www.academia.edu/5717224/Viber_Communication_Security_unscramble_the_scrambled_Contents. [Accessed 25 Jul. 2017].

Bellman, S., Johnson, E.J., Kobrin, S.J., & Lohse, G.L. (2004). International Differences in Information Privacy Concerns: A Global Survey of Consumers. The Information Society, 20, 313-324.

Benenson Z, Reinfelder L. (2013). Should the users be informed? On differences in risk perception between Android and iPhone users. In Symposium on Usable Privacy and Security (SOUPS), Newcastle, UK.

Boksasp, Trond, and Eivind Utnes (2012). "Android apps and permissions: Security and privacy risks."

Buchenscheit, A., Könings, B., Neubert, A., Schaub, F., Schneider, M. and Kargl, F. (2014). Privacy implications of presence sharing in mobile messaging applications. Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia - MUM '14.

Capistrano PE, Chena VJ (2015). Information privacy policies: The effects of policy characteristics and online experience. Comp Standards Interfaces 2015;42:24–31.

Fletcher, K. (2003). Consumer Power and Privacy: The Changing Nature of CRM. International Journal of Advertising, 22, 249-272.

Caruso, D. (1998). The Law and the Internet Beware. Columbia Journalism Review, 37 (1), 57-61.

Gillmor, D. (1998). Violating Privacy is Bad Business. Computerworld, 32 (12), 38-39.

Global Internet User Survey (2012), Internet Society http://www.internetsociety.org/survey

GreekPhone. Lexicon Software GreekPhones. [online]. Available at: http://www.11888.gr. [Accessed 25 Jun. 2017].

Dinev, T. & Hart, P. (2004). Internet Privacy Concerns and Their Antecedents—Measurement Validity and a Regression Model. Behavior and Information Technology, 23 (6), 413-422.

Google (2011). Google Image Search API (Deprecated). [online]. Available at: https://developers.google.com/image-search/v1/jsondevguide. [Accessed 25 Jun. 2017].

Google (2016). What is Google Custom Search. [online]. Available at: https://developers.google.com/custom-search/json-api/v1/overview. [Accessed 25 Jun. 2017].

Govani T., Pashley H. (2007). Student Awareness of the Privacy Implications while Using Facebook. Unpublished manuscript, 2007. [online] Available at :http://lorrie.cranor.org/courses/fa05/tubzhlp.pdf . [Accessed 25 Jun. 2017].

Gross, R., Acquisti, A. and Heinz, H. (2005). Information revelation and privacy in online social networks. Proceedings of the 2005 ACM workshop on Privacy in the electronic society - WPES '05, [online] pp.71-80. Available at: https://www.heinz.cmu.edu/~acquisti/papers/privacy-facebook-gross-acquisti.pdf [Accessed 25 Jun. 2017].

Heather Clancy (2013). Yearning for a unified contact list? Sync.Me wants your number. [online]. Available at: http://www.zdnet.com/article/yearning-for-a-unified-contact-list-sync-me-wants-your-number. [Accessed 25 Jun. 2017].

Loie Favre (2014). HOW TO SYNC CONTACT PHOTOS FROM FACEBOOK. [online]. Available at: https://www.androidpit.com/how-to-sync-contact-photos-from-facebook. [Accessed 25 Jun. 2017].

McDonald A., Cranor L.F., (2008). The Cost of Reading Privacy Policies. I/S: A Journal of Law and Policy for the Information Society 2008. Privacy Year in Review issue. [online] Available at: http://lorrie.cranor.org/pubs/readingPolicyCost-authorDraft.pdf.[Accessed 25 Jun 2017].

Olmstead K. and Atkinson, M. (2015). Apps permissions in the google play. [online] Pew Research Center. Available at: http://www.pewinternet.org/2015/11/10/the-majority-of-smartphone-owners-download-apps/ [Accessed 25 Jun. 2017].

Pew Research Center (2015). Google Play Store Apps Permissions. [online] Available at: http://www.pewinternet.org/interactives/apps-permissions. [Accessed 25 Jun. 2017].

Pitkänen, O. and Tuunainen, V. (2012). Disclosing Personal Data Socially — An Empirical Study on Facebook Users' Privacy Awareness. Journal of Information Privacy and Security, 8(1), pp.3-29.

Rakib A, Ho S. (2011). Privacy concerns of users for location-based mobile personalization, In Proceedings of the International Conference on Information Resources Management;2011.

Rust, R.T., Kannan, P.K., & Peng, N. (2002). The Customer Economics of Internet Privacy. Journal of the Academy of Marketing Science, 30, 455-464.

Statista (2017a). Number of monthly active Facebook Messenger users from April 2014 to April 2017. Available at https://www.statista.com/statistics/417295/facebook-messenger-monthly-active-users . [Accessed 25 Jun 2017].

Statista (2017b). Number of monthly active WhatsApp users worldwide from April 2013 to January 2017. Available at https://www.statista.com/statistics/260819/number-of-monthly-active-whatsapp-users . [Accessed 25 Jun 2017].

Statista (2017c). Number of unique Viber user IDs from June 2011 to March 2017. Available at https://www.statista.com/statistics/316414/viber-messenger-registered-users . [Accessed 25 Jun 2017].

Sync.me. Caller ID and Phone Number Search. [Online]. Available at: https://sync.me. [Accessed 25 Jun. 2017].

Tow W, Dell P, Venable J. (2008). Understanding information disclosure behaviour in Australian Facebook users. J Inform Technol 2008;25(2):126–36. [Online] Available at: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1004&context=acis2008 [Accessed 25 Jun. 2017].

Truecaller: Caller ID & Dialer. [online] Available at: https://play.google.com/store/apps/details?id=com.truecaller. [Accessed 25 Jun. 2017].

TRUSTe (2014), TRUSTe 2014 US Consumer Confidence Privacy Report, Consumer Opinion and Business Impact, Available at: http://download.truste.com/dload.php/?f=4HKV87KT-447 . [Accessed 25 Jun. 2017].