Collateral damage of Facebook Apps: an enhanced privacy scoring model *

Iraklis Symeonidis[1], Filipe Beato[1], Pagona Tsormpatzoudi[2] and Bart Preneel[1]

[1]ESAT/COSIC, KU Leuven and iMinds ,
first.lastname@esat.kuleuven.be
[2]Research Unit ICRI/CIR , pagona.tsormpatzoudi@law.kuleuven.be

Abstract

Establishing friendship relationships on Facebook often entails information sharing which is based on the social trust and implicit contract between users and their friends. In this context, Facebook offers applications (Apps) developed by third party application providers (AppPs), which may grant access to users' personal data via Apps installed by their friends. Such access takes place outside the circle of social trust with the user not being aware whether a friend has installed an App collecting her data. In some cases, one or more AppPs may cluster several Apps and thus gain access to a collection of personal data. As a consequence privacy risks emerge. Previous research has mentioned the need to quantify privacy risks on Online Social Networks (OSNs). Nevertheless, most of the existing works do not focus on the personal data disclosure via Apps. Moreover, the problem of personal data clustering from AppPs has not been studied. In this work we perform a general analysis of the privacy threats stemming from the personal data requested by Apps installed by the user's friends from a technical and legal point of view. In order to assist users, we propose a model and a privacy scoring formula to calculate the amount of personal data that may be exposed to AppPs. Moreover, we propose algorithms that based on clustering, computes the visibility of each personal data to the AppPs.

*Part of this work has been submitted to ICISSP conference 2015

1 Introduction

Facebook has altered the social ecosystem in a remarkable way offering a plethora of easy–to–use tools enabling direct and instant interaction amongst users. Besides such tools, Facebook offers applications (Apps) providing games, lifestyle and entertainment possibilities developed by third parties. Such applications may disclose the user's personal data to a subset of other users of Facebook, the OSN provider himself (Facebook), as well as the third party application providers (AppPs) [**?**, **?**]. In some cases, one or more AppPs can cluster several Apps [**?**], which may eventually grant them access to larger amounts of user's personal data.

As a consequence, personal data disclosure may pose significant risks to users' privacy and has prompted serious concerns among the users, the media and the research community [**?**, **?**, **?**, **?**]. When a user shares information with friends through Facebook, this user relies on the social trust and implicit contract established with her friends in the audience. However, the user who shares personal information has no idea whether a friend has installed Apps that may access the user's information without participating in the circle of social trust and whether her personal data may be furthermore exposed to other AppPs.

From a legal point of view, the above discussion causes data protection law concerns. Data protection law provides a framework for protection of users' fundamental rights, in particular with regards to the right to data protection as well as the right to privacy when it is interfered due to the processing of personal data. Personal data, pursuant to Article 2(a) of the Data Protection Directive 95/46/EC [**?**] "refers to any information relating to an identified or identifiable natural person (data subject)" which in the scope of this paper corresponds to the users measurable profile items processed by the App. In the above discussion, data protection law may be infringed in relation with two aspects: First, data processing lacks legitimacy, as the user has not given her unambiguous consent to the App to process her personal data. Second, data processing lacks transparency, as the user may be totally unaware of the data processing that may take place.

There exists a considerable amount of related work regarding the user's privacy on Facebook, including work that considers installed Apps [**?**, **?**, **?**, **?**]. However, there is currently limited work informing users how their friends Facebook ecosystem affects their privacy.

The fact that Apps are hosted on AppPs beyond Facebook's control doesn't render handling the situation easily. We define as the *collateral damage of Facebook Apps* the privacy issues that arise by: (1) the acquisition

of users' personal data via Apps installed by their friends on Facebook, (2) the clustering of users' personal data via AppPs, exposing these data outside Facebook ecosystem without users' prior knowledge.

Goals and outline Motivated by the above discussion and based on the assumption of privacy as a good practice, illustrated by Diaz and Gürses [**?**], we focus on a solution which, from a legal point of view, implements transparency. We propose a model and a privacy scoring formula, which throws light on how the user's data disclosure takes place via their friends' installed Apps. We thus consider the privacy score of a user as an indicator of her privacy risk. The higher the privacy score the higher the threat for a user. The solution is proposed as a Privacy Enhancing Technology (PET), which is able to raise awareness on personal data (i.e., profile item) collection and may eventually support the user's decisions about personal data sharing [**?**]. The model goes a step beyond and particularly focuses on the case where a set of AppPs gains access to users' profile items via Apps installed by a user's friend. Our solution is in line with the principle of Data Protection by Default, introduced in Article 23 (2) of the proposed Data Protection Regulation which requires mechanisms that "by default ensure that the users are able to control the distribution of their personal data" [**?**]. Data protection by default intends to mitigate privacy risks stemming from users' asymmetrical information, i.e., lack of knowledge or understanding in relation to the data processing that takes place [**?**], as for instance, through their friends' Apps. In the context of this paper, data protection by default solutions, such as the scoring formula which assist privacy management, raise awareness and enhance user empowerment.

The rest of this paper is organized as follows. Section 2 reviews the related work with respect to the privacy issues that Apps introduce on Facebook as well as the need to quantify the user's privacy risk. Section 3 describes the threat model in the case of the *collateral damage of Facebook Apps*.

2 Related Work

The privacy score of a user can be estimated in different ways. Maximilien et al. [**?**] initially proposed a Privacy–as–a–Service (Paas) formula to calculate the privacy score in a social graph, as the product of *sensitivity* and *visibility* of a profile item. Liu and Terzi [**?**] extended this work [**?**] and proposed a framework for computing privacy scores for OSNs, using a probabilistic

model based on the Item Response Theory (IRT). Although, both IRT and PaaS present interesting results, they are not designed for a complex scenario as in the case of the Apps on Facebook [?]. Minkus et al. [?] estimated the sensitivity and visibility of the privacy settings based on a survey of 189 participants. Moreover, Nepali and Wang [?] proposed a privacy index to evaluate the inferring attacks as described by Sweeney and Latanya [?], whereas Ngoc et al. [?] introduced a metric to estimate the potential leakage of private information from public posts in OSNs.

Our work extends the privacy scoring formulas initially introduced in [?, ?] for the case of the *collateral damage of Facebook Apps*. Although, both formulas present interesting results, they work miss to holistically describe the case of the Apps on Facebook and its effects on the user's privacy. These works are mainly based on a privacy score applied on an artificial graph. However, they don't describe the particularities and the privacy effects of the Apps on Facebook; the case of users' profile items and the exposure via their friend Apps is not taken into account.

Moreover, different works demonstrated the privacy issues of the Apps on Facebook, from the collection of user's profile items. Chia et al. [?] showed that certain Apps collect more information than required. Frank et al. [?] revealed the existence of malicious Apps that deviate from the generic permissions pattern acquiring more information from the users. Subsequently, Chaabane et al. [?] identified that Apps gain tracking capabilities, and can later disseminate the collected information to "fourth party" organizations [?] following additional incentives. Similarly, Huber et al. developed an automatic evaluation tool, AppInspect [?], and demonstrated the security and privacy leakages of a large set of Facebook Apps. Moreover, Biczók and Chia [?] described the issue of users' information leaked through their friends via Apps on Facebook. This work introduced a game theory approach to simulate an interdependent privacy scenario of two users and one App game while Pu and Grossklags [?] proposed a formula to estimate the payoffs.

3 Privacy Threats

This section describes the major threats that Apps introduce to user's privacy for the case of the *collateral damage of Facebook Apps*.

Aside from the communication channels already available on Facebook (e.g., likes, posts and personal messages), users can also use Apps for an improved personalized functionalities. These Apps can be downloaded and installed through the Facebook App center [?]. For installation and opera-

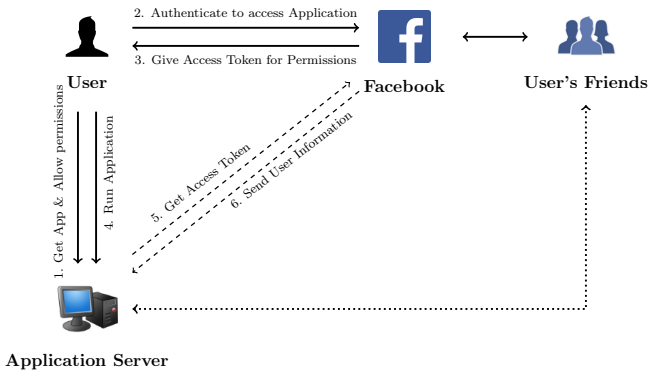

Figure 1: Facebook Applications Architecture Overview.

tion, each App requests from the user a set of *permissions*, granting the App the right to access and collect additional information. This is done by an access token provided by Facebook upon user's authorization (step 1 to 4 in Figure 1). After the user's approval, Apps can collect the user's personal data and store it in servers outside Facebook's ecosystem and user's control (step 5 to 6 in Figure 1).

To better adjust the visibility of the user's personal data, Facebook offers a set of *privacy settings* to its users. It provides four granular levels such as "only me", "friends", "friends to friends" and "public". For the case of Apps on Facebook, the privacy setting "only me" forces the visibility of the personal data only to the user. However, privacy settings of "friends", "friends of friends" and "public" are exposing equally the user's personal data to AppPs via their friend Apps, making them available an external server, outside the Facebook ecosystem.

Even if Apps are available after a thorough review from Facebook, this is not sufficient to prevent users' personal data exposure [**?**]. Due to the server–to–server communication (step 5 to 6 in Figure 1), the offline interaction between Facebook and an AppP makes any protection mechanism hard to apply [**?**, **?**]. As a result, the user's profile items can arbitrary be retrieved by an AppP without any chance of notification or on demand approval by the user.

User's information is exposed by their friends Initially, Facebook provided a set permissions to the AppPs, such as $friends\_birthday$, and $friends\_location$ (Table 1), giving the App the right to access and collect

Permission Group	Permissions	Profile Items
Public profile (default)	public_profile①②	id, name, first_name, last_name, link, gender, locale, timezone, updated_time, verified
App friends	user_friends①②	bio, birthday, education, first_name, last_name, gender, interested_in, languages, location, political, relationship_status, religion, quotes, website, work,
Extended Profile Properties (xpP)*	friends_about_me①, friends_actions①, friends_activities①, friends_birthday① friends_checkins①, friends_education_history①, friends_events①, friends_games_activity①, friends_groups①, friends_hometown①, friends_interests①, friends_likes①, friends_location①, friends_notes①, friends_online_presence①, friends_photo_video_tags①, friends_photos①, friends_questions①, friends_relationship_details①, friends_relationships①, friends_religion_politics①, friends_status①, friends_subscriptions①, friends_website①, friends_work_history①	about_me, actions, activities, birthday checkins, history, events, games_activity, groups, hometown, interests, likes, location, notes, online_presence, photo_video_tags, photos, questions, relationship_details, relationships, religion_politics, status, subscriptions, website, work_history
Extended Permissions (xP)*	read_mailbox①②	inbox

Table 1: Facebook application permissions and the corresponding profile items. Permission availability to API v1.x (①) and v2.x (②)

users' personal data via their friends. Currently, Facebook API version 1.x is obsolete with the $friends\_xxx$ permissions being removed. This set of permissions is still available and being used by Apps [**?**]. The updated API version 2.x replaced the $friends\_xxx$ permissions with the $user\_friends$. Although the newer API had to be in line with the regulations of Federal Trade Commission (FTC) in U.S. [**?**], we identified from our analysis that it disclosures up to fourteen profile items of a user via their friends; maintaining the privacy concerns of collateral damage of Facebook Apps still open and active. A more detailed view on the available permissions of Apps is given in Table **??** and Table **??** (Appendix).

Moreover, the case of Apps requesting permissions through strangers

Permissions	10k Monthly	10k Weekly	10k Daily
Profile Info	100%	100%	100%
Email	67.22%	63.99%	61.82%
Publish	68.99%	67.44%	64.32%
Likes	11.77%	13.37%	13.18%
Location	8.36%	8.33%	8%
Stream	17.93%	18.27%	20.64%
Manage	1.32%	1.46%	1.55%
Friends	10.23%	10.13%	9.77%
Personal Mailbox	1%	1%	1%

Table 2: Facebook top application permissions

(non-friends) who participated in the same group conversation with the user. This is the case, for instance, for the permission *read_mailbox*. Even though *read_mailbox* appears to be used by only 1% of the Apps, as it is illustrated by the table 2), the severity of the risks it may entail for the user is significant. The mere exchange of text messages in a group conversation in which a totally stranger (non-friend) who has installed *read_mailbox*, participates, renders the user's personal data deriving from the conversation accessible to the App. This personal data can be the content of the conversation as well as the time that the communication took place.

Company	Num of Apps	Apps $\geq$ 10k MAU [1]
Vipo Komunikacijos	163	99
Telaxo	136	118
Mindjolt	120	32
superplay!	81	8

Table 3: Third party applications per third party providers

Clustering Third party providers can be owners of several Apps. For instance, there are AppPs with up to 160 Apps for an amount of more than 10k monthly active users, as it is described in Table 3. As a consequence, one or more AppPs may cluster several Apps and thus more profile items. The amount of profile items that can be retrieved are more and equal to the union of all the collected personal data under the same AppP. Moreover, every App retrieves the Facebook's user ID which can lead to uniquely identify a user and accurately correlate the collected personal data of her from each App.

Legal issues From a legal perspective, one of the main challenges to data protection attached to the Apps permissions, as described above, is the fact that personal data processing may lack legitimacy. Article 7 of Data Protection Directive 1995/46/EC [**?**] provides a limited number of legal grounds for data processing. One such ground in order to perform personal data processing is the user's unambiguous consent. As enshrined in Article 7a, the data controller, i.e., Facebook or Apps, may collect, store, use, further disclosed if the user has given her consent. For the consent to be valid, it has to fulfil certain criteria: it has to be given prior to the data processing and in a "free" way and has to be sufficiently specified and informed (Art. 29 WP 2011) [**?**].

In fact, this is not the case with the third party Apps, as they may proceed to personal data processing not with users', but with user friends' consent. In other words, consent may be provided by the actual user of the application and not by the data subject, whose data are going to be processed in the end. On Facebook Apps settings, users allow by default their data to be used by the applications used by their friends without their consent under the title "Apps other use" unless they manually unclick the relevant boxes. One could claim that consent has been theoretically given, however, according it should not be considered as valid as it is not informed. Privacy default settings are such that users are totally unaware of the fact that they have to unclick the boxes in order to prevent such data processing. It is worth to be mentioned that in a relevant case in the U.S., the Federal Trade Commission required that such applications cannot imply consent but rather consent should be affirmatively expressed [**?**].

Further, with regards to the obligation of the data controller (Facebook or Apps) to transparency, it should be noted that in both cases users have no sufficient information about the nature and amount of data that will be collected, the purposes that the data will be used for and the third parties that data may be shared with so in other words data processing goes far beyond the user's legitimate expectations. This interferes with the principle of fairness and transparency stemming from Article 10 of Data Protection Directive 46/1995/EC [**?**]. In relation with the same matter in the U.S., the Federal Trade Commission stressed the need to keep users informed in case any disclosure exceeds the restrictions imposed by the privacy setting(s) of the third party application [12], which can possibly be the case with permissions such as *user_friends*.