Collateral damage of Facebook Apps: an enhanced privacy scoring model

Iraklis Symeonidis, Filipe Beato
ESAT COSIC, KU Leuven and iMinds
first.lastname@esat.kuleuven.be

Pagona Tsormpatzoudi
ICRI/CIR, KU Leuven and iMinds
pagona.tsormpatzoudi@law.kuleuven.be

Bart Preneel
ESAT/COSIC, KU Leuven and iMinds
first.lastname@esat.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 users 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.**

I. 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) [5], [20]. In some cases, one or more AppPs can cluster several Apps [11], 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 [1], [16], [17], [32]. 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 [27] "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 [3], [5], [6], [15]. 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.

a) Goals and outline: Motivated by the above discussion and based on the assumption of privacy as a good practice, illustrated by Diaz and Gürses [7], 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 [24]. 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" [26]. 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 [24], 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 II 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 III describes the threat model in the case of the *collateral damage of Facebook Apps*. Further, section IV provides an overview of the proposed model as well as describes the main components of the privacy score formula, while Section V illustrates the privacy score formula. Lastly, Section VI summarizes and discusses the paper.

II. Related Work

The privacy score of a user can be estimated in different ways. Maximilien et al. [19] 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 [18] extended this work [19] 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 [22]. Minkus et al. [21] estimated the sensitivity and visibility of the privacy settings based on a survey of 189 participants. Moreover, Nepali and Wang [22] proposed a privacy index to evaluate the inferring attacks as described by Sweeney and Latanya [30], whereas Ngoc et al. [23] 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 [18], [19] 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. [6] showed that certain Apps collect more information than required. Frank et al. [13] revealed the existence of malicious Apps that deviate from the generic permissions pattern acquiring more information from the users. Subsequently, Chaabane et al. [5] identified that Apps gain tracking capabilities, and can later disseminate the collected information to "fourth party" organizations [4] following additional incentives. Similarly, Huber et al. developed an automatic evaluation tool, AppInspect [15], and demonstrated the security and privacy leakages of a large set of Facebook Apps. Moreover, Biczók and Chia [3] 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 [28] proposed a formula to estimate the payoffs.

III. 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*.

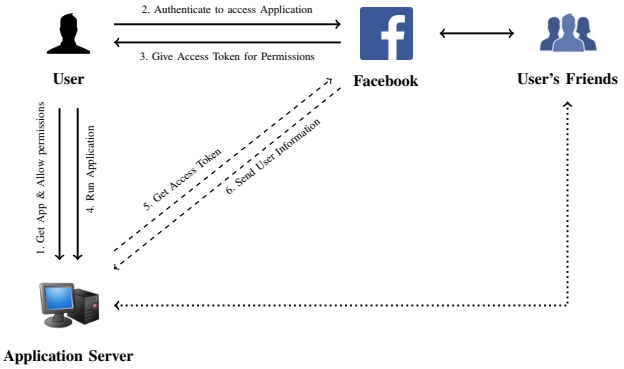

Fig. 1. Facebook Applications Architecture Overview.

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 [9]. For installation and operation, 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).

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

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 [10]. 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 [8], [14]. 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.

b) 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 I), giving the App the right to access and collect 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 [2]. 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. [11], 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 V and Table VI (Appendix).

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 II. FACEBOOK TOP APPLICATION PERMISSIONS

Moreover, the case of Apps requesting permissions through strangers (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 II), 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 III. THIRD PARTY APPLICATIONS PER THIRD PARTY PROVIDERS

c) 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 III. 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.

d) 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 [27] 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) [25].

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 [12].

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 users legitimate expectations. This interferes with the principle of fairness and transparency stemming from Article 10 of Data Protection Directive 46/1995/EC [27]. 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*.

IV. THE COLLATERAL DAMAGE OF FACEBOOK APPS

In this section, we introduce a model to estimate the exposure of user's profile items for the case of the *collateral damage of Facebook Apps*. The necessary notations and the main premises of the privacy scoring (PS) formula are introduced.

A. Notations

When a user's friend installs an App on Facebook, several permissions are requested. Each permission provides access to

a set of profile items. It can correspond to one or multiple profile items, as described in Section III. The $user\_friends$ permission contains up to fourteen profile items by the time this paper was written. We denote as α the permission acquired while installing an App and i the corresponding profile item that is retrieved such that $i = \{birthday\}$ and $\alpha = \{friends\_birthday\}$. Each $i \in \{1, ..., n\}$ and $a \in \{1, ..., A\}$ where n is the total number of profile items and A the total number of permissions available on Facebook's API. The relation of permission–profile item is denoted by i_α such that $i_{friends\_birthday} = \{birthday\}$. However, without loss of generality both notations i and i_α refer to a profile item. Our approach aims to holistically analyse and evaluate the PR of a user for the case of the user's information being exposed via their user's friends Apps outside of Facebook ecosystem.

A user can restrict the visibility of a profile item by adjusting the privacy settings (Section III). Facebook provides four granular levels such as "only me", "friends", "friends to friends" and "public". The first makes the profile item accessible only to the owner, while the last makes it available to every user inside the Facebook ecosystem. However, AppPs on Facebook can collect a user's profile item via her friends Apps outside of Facebook ecosystem for every privacy setting greater or equal to "friends".

	AppP 1		AppP j		AppP N	
Profile item 1	R(1, 2)		R(1, j)		R(1, N)	
Profile item i			$R(i,j)$			$\|R_i\|$
Profile item n	$R(n,1)$				R(n, N)	
			$\|R^j\|$			

TABLE IV. DICHOTOMOUS MATRIX R

To simulate whether a profile item i is visible to the AppPs we use a two dimensional matrix R, as it is illustrated in Table IV-A. The size of R is $n \times N$ where n refers to the number of profile items and N to the number of recipients j. The recipient j is the user's friend and particularly the AppP that collects a profile item i. The rows of R correspond to profile items i, whereas the columns to recipient j. Each $R(i, j)$ stores the availability of a profile item i for the recipient j.

Each profile item i availability depends on the privacy settings of Facebook denoted by ps and the user's decision. For the Apps on Facebook, the privacy setting of $ps_0 = only\ me$ forces the $R(i, j) = 0$. Privacy settings of $ps_1 = friends$, $ps_2 = friends\ of\ friends$ and $ps_3 = public$ are considered to be true (i.e., $R(i, j) = 1$) exposing the same amount of user profile items i to the Apps installed by their friends and thus to the AppPs. As a consequence $R(i, j) \in \{0, 1\}$ and the matrix R is *dichotomous*. For instance, ps_1, ps_2, ps_3 privacy setting in combination with the $\alpha = \{friends\_birthday\}$ permission reveals the $i_\alpha = \{birthday\}$ profile item only the recipient j which $R(i_\alpha, j) = 1$.

B. Compute the contents of R

Data: G: social graph, App_DB: The list of Apps with their profile items i and recipients j (i.e., $AppP$), privacy settings ps
Result: Matrix $R(i, j)$

```
/* Get the number of user's friends*/
F ← number(get_list_of_friends(G));
for f ← 1 to F do
    /*Get the List of user's Friends*/
    LF[f] ← get_list_of_friends(G);
end
for f ← 1 to F do
    /*Get the List of friends Apps*/
    LApps[f] ← get_list_of_friend_Apps(LF[f]);
    for i ← 1 to n do
        /*Retrieve the List of profile items i for each App*/
        LApp_i[f][i] ← query_App_DB(LApps[f]);
    end
end
/*Create the union of profile items per AppP j*/
for j ← 1 to N do
    for i ← 1 to n do
        switch ps do
            case only me
                R(i, j) ← 0;
            case friends, friends of friends, public
                R(i, j) ← Union_of_ij(LApp_i[f][i], j);
        end
    end
end
```

Algorithm 1: RFinder

The contents of the dichotomous matrix R are computed (i.e., RFinder), as it is described by the algorithm 1. The inputs are threefold, it receives: (1) The user's friends from Facebook. Facebook is social graph denoted by G, which contains users as nodes and edges as relationships between vertices (i.e., friendship relation) [29], [31]. (2) The App features from the Appinspect dataset [2] denoted by App_DB, which is continuously extended and publicly available online. It is formed by a total of 25k Facebook Apps containing the respective profile item and the AppP of each App. (3) The privacy settings of a user to indicate whether the collateral damage of Facebook Apps privacy threat is inactive (i.e., is potentially active due to the default privacy settings).

The number of friends F initially are calculated by the $get\_list\_of\_friends(G)$ function, analysing the friendship connections from the graph G. The friends of a user and their installed Apps are populated and stored in matrices $LF[f]$ and $LApps[f]$ using the functions $get\_list\_of\_friends(G)$ and $get\_list\_of\_friends\_Apps(LF[f])$ respectively. For every App the related profile items i are collected querying the App_DB dataset with the $query\_App\_DB(LApps[f])$ function.

The relationship between AppP and App can one to one or one to multiple respectively. For instance, an AppP can own several Apps, as it is described in Section III. As a consequence, the profile items that are collected by an AppP are all the profile items i that are acquired by the user's friends Apps under this AppP. However, similar profile items i can be acquired from multiple Apps. The union of all the collected profile items under the same AppP is performed and estimated by the function $Union\_of\_ij(LApp\_i[f][i], j)$; indicating whether a profile item i is exposed to a particular

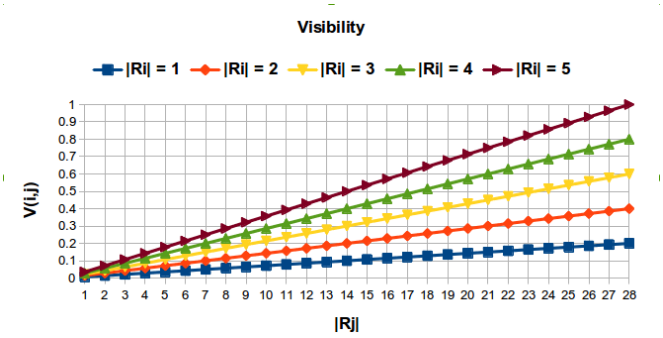

Fig. 2. Graph that demonstrates the V(i, j) values

recipient j (i.e., AppP). For every privacy setting except $ps = onlyme$ the availability of every profile item i to the recipient j identified and stored in the $R(i, j)$ matrix.

C. Visibility

The visibility of a profile item i captures how much i is known to the recipient j, where the more it spreads the higher the visibility. For instance, a profile item i has higher visibility when $R(i, j) = 1$ rather than $R(i, j) = 0$.

It is considered the visibility of a profile item i for a recipient j to be a random variable described by a probability distribution. For availability of a profile item i in dichotomous matrix R it is denoted P_{ij} the probability that a recipient j has $R(i, j) = 1$ with $P_{ij} = Prob\{R(i, j) = 1\}$. Liu and Terzi [18] described that the visibility of a profile item i to a recipient j is $V(i, j) = P_{ij} \times 1 + (1 - P_{ij}) \times 0 = P_{ij}$.

Definition The visibility is denoted by $V(i, j)$ for each profile item $i \in \{1, ..., n\}$ and recipient $j \in \{1, ..., N\}$. Assuming independence of profile items i and recipients j the probability of P_{ij} is the product of 1 in row R_i (i.e., $\frac{|R_i|}{N}$) times the probability of 1 in column j (i.e., $\frac{|R^j|}{n}$) [18].

$$V(i, j) = P_{ij} = \frac{|R_i|}{N} \times \frac{|R^j|}{n} \tag{1}$$

Intuitively, the higher the visibility $V(i, j)$ the more a recipient j (i.e., AppP) collects multiple profile items (i.e., $|R^j|$) and the more a profile item i is exposed (i.e., $|R_i|$). The visibility $V(i, j)$ is monotonically increasing, as it is illustrated by the Figure 2. For instance, for a profile item i that is exposed $|R_i| = 5$ number of times out of $N = 10$ and a recipient j that acquires $|R_j| = 4$ profile items out $n = 28$ the visibility of a profile item i been exposed from a user's friend App to a recipient j (i.e., AppP) is $V(i, j) = \frac{5}{10} \times \frac{4}{28}$ (0.071). Whereas for the case of $|R_i| = 5$ and $|R_j| = 10$ the $V(i, j) = \frac{5}{10} \times \frac{10}{28}$ (0.178)

D. Sensitivity

A profile item i can be more sensitive than another and thus the sensitivity of i depends on the profile item itself. Exposing personal messages (i.e., $i_{read\_mailbox} = \{inbox\}$) and the user's birthday (i.e., $i_{friends\_birthday} = \{birthday\}$) should have a different impact for the user, being the first more sensitive than the second.

Definition The sensitivity of a profile item $i \in \{1, ..., n\}$ is denoted by β_i and depends on the characteristic of the item i itself. Is the ratio of users that do not expose a profile item i to the whole set or recipients N.

$$\beta_i = \frac{N - |R_i|}{N} \tag{2}$$

Intuitively, the higher sensitive β_i of a profile item i the less people discloses i. The sensitivity β_i monotonically decreases, as it is described by the equation 2. For instance, for a profile item i that is exposed $|R_i| = 5$ number of times out of $N = 10$ the $\beta_i = \frac{10-5}{10}$ (0.5). Whereas for the case of $|R_i| = 7$ the $b_i = \frac{10-7}{10}$ (0.3).

V. Privacy Score

This section describes the proposed Privacy Score (PS) formula, for the case of the *collateral damage of Facebook Apps*. The PS of a user is an indicator of her privacy risk. The higher the privacy score the higher the threat for a user.

Liu and Terzi [18] proposed the matrix R to represent the potential availability of user's profile items to the social graph G. However, this doesn't represents the case of profile items being exposed to several third party applications (Apps) and to the third party application providers (AppPs). For the case of the collateral damage of Facebook Apps, every profile item i is exposed to a user's friend, to her installed Apps and finally to the AppPs. Without losing the functionality of R, each content $R(i, j)$ is referred to the exposure of a profile item i to the AppPs.

A. Privacy Score of a user

The Privacy Score (PS) of a user is related to the *visibility* and *sensitivity* of each profile item i being exposed to several recipients j. Profile items that are acquired by several user's friends Apps can be exposed to a smaller set of AppPs. The user's PS is estimated by the summation of all $PS(j)$ that each recipient j (i.e., AppP) introduces.

Definition The privacy score is denoted by PS for each profile item $i \in \{1, ..., n\}$ and recipient $j \in \{1, ..., N\}$. This score presents the summation of each privacy score $PS(j)$ every recipient j introduces, as the privacy effect of profile items i being collected by AppPs j. Each $PS(j)$ is the product of *sensitivity* β_i and *visibility* $V(i, j)$.

$$PS = \sum_{j=1}^{N} \sum_{i=1}^{n} PS(j) = \sum_{j=1}^{N} \sum_{i=1}^{n} \beta_i \times V(i, j) \tag{3}$$

The PS monotonically increases, with the higher the values of privacy score the more exposed a user is (i.e., profile items), as it is described by the equation 3.

B. Privacy Score of a user's friend

Data: G: social graph, App_DB: The list of Apps with their profile items i and recipients j (i.e., $AppP$), privacy settings ps
Result: Matrix $R(i,j)$
$f \leftarrow friend$;
/*Get the List of friend Apps*/
$LApps[f] \leftarrow get\_list\_of\_friend\_Apps(LF[f])$;
for $i \leftarrow 1$ **to** n **do**
  /*Retrieve the List of profile items i for each App*/
  $LApp\_i[f][i] \leftarrow query\_App\_DB(LApps[f])$;
end
/*Create the union of profile items per AppP j*/
for $j \leftarrow 1$ **to** N **do**
  for $i \leftarrow 1$ **to** n **do**
    switch ps **do**
      case $only\ me$
        $R(i,j) \leftarrow 0$;
      case $friends,\ friends\ of\ friends,\ public$
        $R(i,j) \leftarrow Union\_of\_ij(LApp\_i[f][i], j)$;
    end
  end
end

Algorithm 2: RFinder_f

To calculate the Privacy Score (PS) of a particular friend of a user, the dichotomous matrix $R(i,j)$ should be recomputed. We propose the RFinder_f to identify and cluster the actual exposure of profile items to the AppPs, as it is described by the algorithm 2. The user's friend privacy score can be computed by the equation 3, after the matrix $R(i,j)$ populated.

VI. Summary

Departing from the privacy issues that arise upon installation of third party applications (Apps) via the user's friends on Facebook, this paper proposes a PET solution, which illustrates how the user's data disclosure takes place via their friends' installed Apps. To that end, we verified that several applications collect permissions for considered sensitive information, such as email (68.99%), friends information (10.23%), and, more invasive, private mailbox privileges (1%). We used the publicly available dataset [15] with $25k$ users' Apps and analysed the ones with more than $10\,000$ active users. Whereas the permissions affecting friends data seem limited, the total lack of transparency and opt-out option (lack of valid consent) for the user is quite worrisome.

Our model focused on the privacy impact that arise by the acquisition of users' personal data via Apps installed by their friends on Facebook and the clustering of users' personal data via AppPs, exposing these data outside Facebook ecosystem without users' prior knowledge (i.e., *collateral damage of Facebook Apps*). In order to assist users we proposed a model and algorithms to calculate privacy scores of the user's friends on Facebook based on the profile items which also may be clustered under AppPs. Privacy scores calculate the exposure of user's profile items measuring its sensitivity as well as its visibility to the AppPs. Being able to raise awareness on profile item collection, it is in line with the legal principle of data protection by default as it can potentially support decisions and foster user control on personal data disclosure.

References

[1] Facebook and your privacy: Who sees the data you share on the biggest social network? *http://bit.ly/11WhqWt*. Accessed on Sept. 6, 2012.

[2] AppInspect. Appinspect: A framework for automated security and privacy analysis of osn application ecosystems. http://ai.sba-research.org/.

[3] G. Biczók and P. H. Chia. Interdependent privacy: Let me share your data. In *Financial Cryptography and Data Security - 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers*, pages 338–353, 2013.

[4] A. Chaabane, Y. Ding, R. Dey, M. Ali Kaafar, and K. Ross. A Closer Look at Third-Party OSN Applications: Are They Leaking Your Personal Information? In *Passive and Active Measurement conference (2014)*, Los Angeles, États-Unis, Mar. 2014. Springer.

[5] A. Chaabane, M. A. Kaafar, and R. Boreli. Big friend is watching you: Analyzing online social networks tracking capabilities. In *WOSN*, WOSN '12, pages 7–12, New York, NY, USA, 2012. ACM.

[6] P. H. Chia, Y. Yamamoto, and N. Asokan. Is this app safe? a large scale study on application permissions and risk signals. In *WWW*, Lyon, France, 04/2012 2012. ACM.

[7] C. Diaz and S. Gürses. Understanding the landscape of privacy technologies. *Proc. of the Information Security Summit*, pages 58–63, 2012.

[8] W. Enck, P. Gilbert, B. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. Sheth. Taintdroid: an information flow tracking system for real-time privacy monitoring on smartphones. *Commun. ACM*, 57(3):99–106, 2014.

[9] Facebook. Facebook application center. https://www.facebook.com/appcenter. Accessed December 13, 2014.

[10] Facebook. Facebook application center guidelines. https://developers.facebook.com/docs/games/appcenter/guidelines. Accessed December 13, 2014.

[11] Facebook. Facebook privacy settings and 3rd parties. http://www.ftc.gov/enforcement/cases-proceedings/092-3184/facebook-inc. Accessed February 08, 2015.

[12] Facebook. Ftc and facebook agreement for 3rd parties wrt privacy settings. http://www.ftc.gov/sites/default/files/documents/cases/2011/11/111129facebookagree.pdf. Accessed February 08, 2015.

[13] M. Frank, B. Dong, A. Felt, and D. Song. Mining permission request patterns from android and facebook applications. In *ICDM*, pages 870–875, Dec 2012.

[14] P. Guard. Cyanogenmod; a privacy guard manager arrives! http://www.mono-live.com/2013/07/cyanogenmod-privacy-guard-manager.html.

[15] M. Huber, M. Mulazzani, S. Schrittwieser, and E. R. Weippl. Appinspect: large-scale evaluation of social networking apps. In *Conference on Online Social Networks, COSN'13, Boston, MA, USA, October 7-8, 2013*, pages 143–154, 2013.

[16] B. Krishnamurthy and C. E. Wills. Characterizing privacy in online social networks. In *WOSN*, WOSN '08, pages 37–42, New York, NY, USA, 2008. ACM.

[17] E. Lalas, A. Papathanasiou, and C. Lambrinoudakis. Privacy and traceability in social networking sites. In *Informatics (PCI), 2012 16th Panhellenic Conference on*, pages 127–132, 2012.

[18] K. Liu and E. Terzi. A framework for computing the privacy scores of users in online social networks. *TKDD*, 5(1):6, 2010.

[19] E. M. Maximilien, T. Grandison, K. Liu, T. Sun, D. Richardson, and S. Guo. Enabling privacy as a fundamental construct for social networks. In *Proceedings IEEE CSE'09, 12th IEEE International Conference on Computational Science and Engineering, August 29-31, 2009, Vancouver, BC, Canada*, pages 1015–1020, 2009.

[20] C. McCarthy. Understanding what Facebook apps really know (FAQ). http://cnet.co/1k85Fys. Accessed Apr. 9, 2014.

[21] T. Minkus and N. Memon. On a scale from 1 to 10, how private are you? scoring facebook privacy settings. In *Proceedings of the Workshop on Usable Security (USEC 2014). Internet Society*, 2014.

[22] R. K. Nepali and Y. Wang. SONET: A social network model for privacy monitoring and ranking. In *33rd International Conference on*

Distributed Computing Systems Workshops (ICDCS 2013 Workshops), Philadelphia, PA, USA, 8-11 July, 2013, pages 162–166, 2013.

[23] T. H. Ngoc, I. Echizen, K. Komei, and H. Yoshiura. New approach to quantification of privacy on social network sites. In *Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on*, pages 556–564. IEEE, 2010.

[24] T. P. C. F. . L. M. D. Notario McDonnel, Troncoso C. "deliverable 5.1 : State-of-play: Current practices and solutions." fp7 pripare project. http://pripareproject.eu/wp-content/uploads/2013/11/D5.1.pdf. Accessed May 04, 2015.

[25] E. Parliament. 01197/11/en wp187. http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp187_en.pdf. Accessed May 04, 2015.

[26] E. Parliament. European parliament legislative resolution of 12 march 2014 on the proposal for a regulation. http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0212+0+DOC+XML+V0//EN. Accessed May 04, 2015.

[27] E. Parliament and of the Council. Directive 95/46/ec of the european parliament and of the council. http://ec.europa.eu/justice/policies/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf. Accessed April 15, 2015.

[28] Y. Pu and J. Grossklags. An economic model and simulation results of app adoption decisions on networks with interdependent privacy consequences. In *Decision and Game Theory for Security - 5th International Conference, GameSec 2014, Los Angeles, CA, USA, November 6-7, 2014. Proceedings*, pages 246–265, 2014.

[29] J. Scott. Social network analysis, overview of. In *Encyclopedia of Complexity and Systems Science*, pages 8265–8279. 2009.

[30] L. Sweeney. Uniqueness of simple demographics in the us population. Technical report, Technical report, Carnegie Mellon University, 2000.

[31] J. Ugander, B. Karrer, L. Backstrom, and C. Marlow. The anatomy of the facebook social graph. *CoRR*, abs/1111.4503, 2011.

[32] Y. Wang, S. Komanduri, P. Leon, G. Norcie, A. Acquisti, and L. Cranor. I regretted the minute I pressed share: A qualitative study of regrets on Facebook. In *SOUPS*, 2011.

APPENDIX

Table V and Table VI illustrates the permissions availiable for the API v1.x and 2.x respectively.

Basic Info (default)	**Extended Profile Properties (xpp)**		**Extended Permissions (xp)**
	User Data	Friends Data	
uid	user_about_me	friends_about_me	ads_management
name	user_actions.books	friends_actions.books	ads_read
first_name	user_actions.music	friends_actions.music	create_event
last_name	user_actions.news	friends_actions.news	create_note
link	user_actions.video	friends_actions.video	email
username	user_activities	friends_activities	export_stream
gender	user_birthday	friends_birthday	manage_friendlists
locale	user_checkins	friends_checkins	manage_notifications
age_range	user_education_history	friends_education_history	manage_pages
	user_events	friends_events	photo_upload
	user_friends	friends_games_activity	publish_actions
	user_games_activity	friends_groups	publish_checkins
	user_groups	friends_hometown	publish_stream
	user_hometown	friends_interests	read_friendlists
	user_interests	friends_likes	read_insights
	user_likes	friends_location	read_mailbox
	user_location	friends_notes	read_page_mailboxes
	user_notes	friends_online_presence	read_requests
	user_online_presence	friends_photo_video_tags	read_stream
	user_photo_video_tags	friends_photos	rsvp_event
	user_photos	friends_questions	share_item
	user_questions	friends_relationship_details	sms
	user_relationship_details	friends_relationships	status_update
	user_relationships	friends_religion_politics	video_upload
	user_religion_politics	friends_status	xmpp_login
	user_status	friends_subscriptions	
	user_videos	friends_website	
	user_website	friends_work_history	
	user_work_history		

TABLE V. FACEBOOK API V1.X

Basic Info (default)	**Extended Profile Properties (xpp)**		**Extended Permissions (xp)**
	User Data	Friends Data	
uid	user_about_me	user_friends	ads_management
name	user_actions.books		ads_read
first_name	user_actions.music		create_event
last_name	user_actions.news		create_note
link	user_actions.video		email
username	user_activities		export_stream
gender	user_birthday		manage_friendlists
locale	user_checkins		manage_notifications
age_range	user_education_history		manage_pages
	user_events		photo_upload
	user_friends		publish_actions
	user_games_activity		publish_checkins
	user_groups		publish_stream
	user_hometown		read_friendlists
	user_interests		read_insights
	user_likes		read_mailbox
	user_location		read_page_mailboxes
	user_notes		read_requests
	user_online_presence		read_stream
	user_photo_video_tags		rsvp_event
	user_photos		share_item
	user_questions		sms
	user_relationship_details		status_update
	user_relationships		video_upload
	user_religion_politics		xmpp_login
	user_status		
	user_videos		
	user_website		
	user_work_history		

TABLE VI. FACEBOOK API V2.X