Evaluating Privacy Risks in Social Networks from the User's Perspective

Michal Sramka

Abstract Determining privacy risks when publishing information on social networks often presents a challenge for the users. A measure of how much of sensitive information users shared with others on a social network website would help the users to understand whether they individually share too much. We survey existing measures that evaluate privacy from the user's perspective or help the user with the privacy risks and related decisions in social networks. We present the Privacy Scores—a measurement of how much sensitive information a user made available for others on a social network website, discuss some of their shortcomings, and discuss research directions for their extensions. In particular, we present our proposal for an extension that takes the privacy score metric from a single social network closed system to include auxiliary background knowledge. Our examples and experimental results show the need to include publicly available background knowledge in the computation of privacy scores in order to get scores that reflect the privacy risks of the users more truthfully. We add background knowledge about users by means of combining several social networks together or by using simple web search for detecting publicly known information about the evaluated users. This is a revision and extension of our former paper.

1 Introduction

Recently there was an explosion of popularity of web sites that allow users to share information. These sites—social-network sites, blogs, and forums such as Google+, Facebook, LinkedIn and others—attract millions of users. The users publish and share information about themselves by creating online profiles, posting blogs and comments. Such information usually contains personal details. Often the users are

M. Sramka (✉)
Faculty of Electrical Engineering and Information Technology, Institute of Computer Science and Mathematics, Slovak University of Technology, Ilkovičova 3, 812 17 Bratislava, Slovakia
e-mail: sramka@stuba.sk

G. Navarro-Arribas and V. Torra (eds.), *Advanced Research in Data Privacy*,
Studies in Computational Intelligence 567, DOI 10.1007/978-3-319-09885-2_14

unaware of the potential risks involved when they are sharing sensitive information online. Quantifying the individuals' privacy risk due to these information-sharing activities of the individuals is a challenging task. Yet the users should know where they stand on the privacy measuring scale.

Securing individuals' privacy in such environments and protecting users against threats such as *identity theft*. *Digital stalking* or *cyberstalking* becomes an increasingly important issue. Both users and service providers recognize the need for users' privacy. The sites may provide some privacy controls. However, the users are faced with too many options and too many controls, and lack the understanding of privacy risks and threats or are unable to accurately asses them. This all contributes to the confusion for the average users, and often results in skipping the complicated and time-consuming tasks of setting the privacy controls that should protect them.

It needs to be noted that there are research directions that try to help the social network users by enabling them to set and personalize their online privacy preferences automatically [1]. But even with properly configured privacy settings for a user profile, some privacy concerns remain. Take for example discussion forums, where tenths or hundreds contributions to multiple discussions of various topics are written by a user. Although the user is careful not to disclose any personally identifiable information in his/her individual posts, personal, sensitive, and private information may be inferred and disclosed by looking at the set of all posts by the user. From the cumulative set of all posts, it may be then possible to profile the user and infer the user's opinions or even identity.

There are primarily two privacy issues [2] in social networks. A lot of research exists dealing with the privacy concerns of publishing the social network data without revealing the identity of an individual. The other privacy risk in social networks comes from the information that has been shared by the users on their profiles:

- *Relationship privacy.* Generally, a social network consist of users and relationships among them. The relationships can be of different kinds—such as "colleague of", "friend of", etc.—and of different trust level—for example, direct relationship, friend-of-a-friend. The availability of information on relationships raises privacy concerns: Knowing who is trusted by whom and to what extent discloses information about the users, their thoughts and feelings. Sometimes just the fact that a relationship exists can be a privacy leakage.
- *Content privacy.* The information content a user shares or publishes on a social network clearly affects the user's privacy. The user can share sensitive or personal information with his/her friends, their friends, or using similar schemed up to sharing completely, that is, basically publishing the information for all. Often some information about the user that s/he wants to keep private can be inferred from other shared information or from information shared with other users.

For a survey of privacy research in social networks see [3], more references are in Sect. 5. Here, we are concerned with the privacy risks from the user's perspective. That is, we focus on measuring the privacy of social network users and helping and enabling them to make informed decisions about their sharing activities, following the research direction of [1, 4–6].

2 Privacy from the User's Point of View

We focus on privacy from the user's point of view. We survey some existing models and measures of user's privacy that empower the users by providing immediate decision support about their actions and their impact on the user's privacy.

Orthogonal to the measures are the tools that help the social network users make informed and wise choices about their privacy settings. We also briefly describe some of these tools.

2.1 Privacy Scores

Privacy Scores by Liu and Terzi [4, 5] were proposed to quantify the privacy risks of individuals posed by their profiles in a social-network site. Focus here is on privacy risks from the individuals' perspective. In the proposed framework, each user in a social network is assigned a privacy score based on the information in his/her profile compared to all available information in all profiles. The score then measures the user's potential privacy risk due to having his/her profile available on the social-network site.

The main drawbacks of this proposal of privacy scores are the concentration only on users' profiles and inconsideration of other publicly available information about the users on the same social network and beyond it. In particular, *background knowledge* about a user is not included in the computation of the privacy score. Background knowledge is some information about an individual that by itself is not a privacy disclosure, but combined with other information it becomes one. Background knowledge is sometimes referred to as external knowledge or auxiliary information.

The value of a Privacy Score is a combination of each one of user's profile items, labelled $1, \ldots, n$, for example, real name, email, hometown, land line number, cell phone number, relationship status, IM screen name, etc. The contribution of each profile item to privacy score is based on sensitivity and visibility. The *sensitivity* β_i depends on the item i itself—the more sensitive the item is, the higher is the privacy risk of it being revealed. The *visibility* of an item i belonging to a user j is denoted $V(i, j)$ and captures how far this item is known in the network—the wider the spread in the network, the higher the visibility.

The privacy score of an item i belonging to a user j is simply $\text{PR}(i, j) = \beta_i \times V(i, j)$. The overall *privacy score* for a user j with n items is then computed as

$$\text{PR}(j) = \sum_{i=1}^{n} \text{PR}(i, j) = \sum_{i=1}^{n} \beta_i \times V(i, j) \ . \tag{1}$$

To keep the privacy score PR a non-decreasing function, in order for it to be a nicely behaving score, both the sensitivity β_i and visibility $V(i, j)$ must be non-negative functions. In practice, the sensitivity and visibility are determined from an $n \times m$

matrix R that represents n items for m users of a single social network. The value of each cell $R(i, j)$ describes the willingness of the user j to disclose the item i. In the simplest case, the value of $R(i, j)$ is 0 if the user j made the item i private and 1 if the item i is made publicly available. From this, the (observed) visibility can be defined as $V(i, j) = R(i, j)$. In a more granular approach, the matrix R can be defined by $R(i, j) = k$, representing that the user j disclosed the item i to all the users that are at most k jumps away in the graph of the social network. Regarding the sensitivity of an item, β_i can be computed using Item Response Theory (IRT) [4, 5]. The IRT can be also used to compute the true visibility of an item for a user.

The privacy score is computed for each user individually. It is an indicator of the user's potential privacy risk—the higher the score of a user, the higher the threat to his/her privacy.

2.2 Privacy Quotient and Privacy Armor

One extension of Privacy Scores comes under the name of *Privacy Quotient* [7]. The authors, similar to our past research [8], have realized that unstructured data pose a problem for privacy score evaluation. The focus here is to evaluate a user's privacy risks in exchanging, sharing, publishing, and disclosing unstructured data—namely, text messages.

A (text) message may contain sensitive information about the user. The message is first checked for any sensitive information such as the user's phone numbers, address, email, or location. The message is then classified as sensitive or non-sensitive by means of a naive binary classifier.

Each sensitive part of the message is treated as an "item" that has some sensitivity. The Privacy Quotient computation is then the same as for the privacy scores, that is, using the Eq. (1). In addition, the message's privacy leakage ϑ is computed as the ratio of sum of all the sensitive parts(items) sensitivities β_i to the sum of all the sensitivities. This privacy leakage ϑ is similar to the computation of the Privacy Index PIDX discussed next in Sect. 2.4 and the Eq. (3).

The authors of Privacy Quotient also proposed the *Privacy Armor* model: For any message a user wants to share with his/her group of friends, the quotient (=score) is computed for not just the message, but an average quotient is computed for the whole group of friends. If the average quotient of the group is above some threshold—some fixed desired quotient value, an alert containing the message's privacy leak may be present to the user, and the message may be anonymized before being sent to the group.

2.3 Privacy-Functionality Score

An interesting research direction motivated by the Privacy Score is the Privacy-Functionality Score [2]. A utility function based on the original privacy scores is

proposed. The utility function measures the rational benefit derived by a user from his/her participation in a social network, in the terms of information acquired versus information provided. The utility is defined as the functionality the user gets divided by privacy risk score the user incurs, that is, the amount of information the user can see about other users in the social network divided by the amount of information the user reveals about himself/herself. The Privacy-Functionality Score of user j, using the notations from Sect. 2.1 and the Eq. (1), is

$$\mathrm{PRF}(j) = \frac{\sum_{j'=1, j' \neq j}^{n} \mathrm{PR}(j')}{1 + \mathrm{PR}(j)} = \frac{\sum_{j'=1, j' \neq j}^{n} \sum_{i=1}^{n} \beta_i \times V(i, j')}{1 + \sum_{i=1}^{n} \beta_i \times V(i, j)} \ . \tag{2}$$

Using this score and considering the social network and privacy to be a game where users are players, the author was able to derive two results.

The first result is when users of a social network try to selfishly maximize this utility score—that is, the users are "free riding" the social network by offering and sharing no information about themselves and only acquiring information from other users. If each user of the social network is independently choosing this strategy, then this case results in the non-functionality and shutdown of the social network.

The second result is based on a game where users choose correlated strategies to jointly get the maximum utility score from the social network. Such strategy indeed exists—the simplest one being "tit-for-tat", where items are disclosed among users sequentially: A user starts the round by revealing the least sensitive item i that has not been shared yet. A next round, where more sensitive items are disclosed, does not start unless all users in a group or the whole social network have revealed the item i. This strategy or a similar reputation-based strategy [2] can be used to assist users in making rational decisions regarding which of his/her attributes the user reveals to other users in a social network.

2.4 PIDX

Privacy Index [9], PIDX, is used to describe a user's privacy exposure factor based on the known (published/shared, in our terminology) attributes. Higher PIDX indicates higher exposure of privacy. PIDX as the proposed privacy risk indicator can be calculated in real time and the value can be used for privacy monitoring and risk control, same as is the case with the previously discussed metrics.

PIDX is defined as the ratio of the sum of the privacy impact factors of the published items, set K, to the sum of the the privacy impact factor of all the items, set I. That is,

$$\mathrm{PIDX} = \frac{\sum_{k \in K} s_k}{\sum_{i \in I} s_i} \times 100 \ , \tag{3}$$

where s_i are privacy impact factors of an item i defined as the sensitivity of the item i, assuming the visibility of the published items to be 1. Since $K \subseteq I$, it is

obvious that Privacy Index PIDX is a score between 0 and 100 and reflects how much sensitive information has the user published. In this sense, the Privacy Index PIDX computation for the user is the same as for the computation of a messages's privacy leakage ϑ of Privacy Quotient, discussed in Sect. 2.2, because sensitivity of an item i is $s_i = \beta_i$.

The authors use the privacy index in a model for privacy ranking and monitoring that employs web searching to look for already known and published items from a user. The web searching can use standard search engines as well as it can be based on the deep web search engines. This is similar to our approach [8].

2.5 PrivAware

Although not a score or metric, PrivAware [6] is a tool to detect and report unintended information loss in social networks. The authors propose to quantify and reduce privacy risks attributed to friends in online social networks. PrivAware tool provides two functions—inference detection and inference reduction.

First, PrivAware infers the attributes (items) of a user based on those of his/her friends. In particular, the tool tries to detect whether attributes of the user at hand can be inferred given all the attributes of his/her friends. PrivAware derives inferences for the following attributes: age, country, state, zip, high school name, high school grad year, university, degree, employer, affiliation, relationship status, and political view.

Second, PrivAware suggests how to change the members of the user's friends to reduce the number of inferable attributes to an acceptable level. The user can simply, but drastically, cut the relationships to his/her friends in order to remove the inferences, or the user can configure his/her privacy settings for these friends in more stringent manner.

2.6 Privometer

The authors in [10] develop a privacy-protection tool, Privometer, to measure the amount of sensitive information leakage in a user's profile. The leakage is indicated by a numerical value. The tool can suggest self-sanitization actions based on the numerical value.

The importance of the research this tool introduced is in looking beyond the publicly available information that the user shares on his/her profile. The model of Privometer also considers substantially more information that a potentially malicious application installed in the user's friend realm can access. Of course, this only applies to social network websites, such as Facebook, that allows applications to access users' information.

2.7 Tools for Social Network Privacy Settings Configuration

As discussed, the social network websites usually provide some privacy controls in the form of a settings page. However, the users are faced with too many options and too many controls, and lack the understanding of privacy risks and threats or are unable to accurately asses them. This all contributes to the confusion for the average users, and often results in skipping the complicated and time-consuming tasks of setting the privacy controls that should protect them.

Here we briefly describe some of the research tools that help the social network users make informed and wise choices about their privacy settings.

2.7.1 Privacy Wizard

Considering this problem of privacy settings, the authors of [1] have proposed a template of a social networking privacy wizard. The idea of the Privacy Wizard is that from a set of user's privacy choices in the form of rules, it is possible to design and build a machine learning model. Such model can then be used to configure the user's privacy settings automatically.

2.7.2 PViz

Another tool for configuring the user's privacy settings is PViz [11]. PViz tool is centered on a graphical display of the privacy choices. It allows the user to understand the visibility of his/her profile according to natural sub-groupings of friends, and at different levels of granularity.

3 Assessing Privacy Risks Beyond Social Networks

Here follows our contribution to the area of assessing and evaluating privacy risks of users in social networks pertained from publishing possibly sensitive information about themselves. This part follows our original research [8].

3.1 Our Contribution

We propose a new concept for Privacy Scores that were introduced in Sect. 2.1. We explore the idea of presenting users with a new privacy score that measures their overall potential privacy risk due to available public information about them. Compared with the original Privacy Scores by Liu and Terzi [4], we overcome the

drawbacks of concentrating only on users' profiles in a single social network, and we include publicly available background knowledge in computation of the new privacy scores. Our new privacy scores metric better represents the potential privacy risks of users and thus helps them make better decision in managing their privacy.

Our results are twofold. Firstly, in Sect. 3.2 we discuss the shortcomings of the privacy scores. We present several opportunities for extending the original privacy scores. With the extension of including background knowledge in mind, we identify some background knowledge that is publicly available but that cannot be easily extracted by computers in an automatic manner. Secondly, we proposed an extension of the privacy score metric that takes it from a closed system evaluating privacy over a single social network to a metric that includes information about the users that comes from outside the social network. In Sect. 3.3, we present examples and experimental results showing paradoxes that may happen when the computation is over only a single social network. Next, in Sect. 4, we extend the computation of privacy scores to include two or multiple social networks. Our final proposal, in Sect. 4.2, uses web searches to include all available public indexed human knowledge in the computation of the privacy score of a user. Thus, our new privacy score reflects the privacy risks of combining user's profile information with available knowledge about the user represented by the web.

Our proposed method for making web search inferences while scoring the privacy risks of individuals can also be seen as a privacy attack. However, we do not explore this direction, as there are already too many attacks, some of them referenced later in Sect. 5. Our contribution rather focuses on helping users achieve their privacy needs and lower their privacy risks. The extended privacy score helps the users to make more informed decisions about their online activities.

3.2 Shortcomings and Opportunities of Privacy Scores

The privacy score, presented in Sect. 2.1, is no doubt a useful metric for each and every user of a social network. Nevertheless, there exist several shortcomings of the originally proposed privacy scores. We list a few of them here. Some of these were already noticed and identified by the authors of the privacy scores, others are just observations, and some are our proposals for further exploration, research, and enhancements of privacy scores.

Regarding the items of a user profile, one can immediately notice hardship in quantifying the items themselves:

- The granularity of profile items is of particular concern. For example, the profile item "personal hobbies" can cover a range of non-private and private information and so its true sensitivity cannot be really established for the general case required by the privacy scores.
- Different profile items have different life-cycles. Some profile items may have a time attribute attached to them—for example, a cell/mobile phone number

or an address are temporary information, while the date of birth or the mother's maiden name are permanent for life. The proposed privacy score, as defined, ignores these facts. We believe that implicit time relevance should be taken into account for more precise evaluation of a user's privacy.

- Impossibility or hardness of including all, possibly private, information in privacy score computation. For example, consider photos: It may be hard or impossible in some cases to (automatically or even by a human involvement/assessment) establish relationships from photos. Or whether a person is drinking alcohol in a photo. Another example are discussion forums: Information is exhibited in natural language form. Determining a political orientation of a user from a single post may not be possible, yet looking at the cumulative set of the user's posts, private information can be inferred about the user (see Sect. 3.3).

Of more concern and interest is the definition, computation and use of sensitivity β_i for item i. As proposed in Privacy Scores, the sensitivity is computed from the matrix R, that is, the sensitivity is based only on the users and items in the single social network. When considering a single social network represented by a matrix R, it is easy to get a wrong perception of privacy due to the limited information about the users.

- The sensitivity β_i computed for an item i would reflect the true real-world sensitivity of this item only if the distribution of people in the social network would mirror the real-world distribution. Obviously, many social networks are not like this, and so paradoxes are likely because of this fact. For example, take a date of birth that most people consider a sensitive and private information. However, if everybody in a social network reveals his/her date of birth, then this item will be considered as not sensitive at all (because everybody reveals it). Paradoxes on the other side of the spectrum are possible, too. For example, if an item in a social network is filled only by one or a few users, because the other users are too lazy to fill it in, then the item will be considered sensitive (by the computation of sensitivity), although the item is far from being considered sensitive or private in the real life. For this reasons, the definition of the sensitivity is not the best possible as it does consider only published information and not the true perception of privacy of the users.
- No background knowledge inclusion, and so no inference detection or control: A privacy metric should include "background knowledge" (auxiliary information or external knowledge) in establishing a score for a user. Speaking more generally, a single social network or any closed system evaluation is not sufficient for real and proper privacy evaluation of a user.
  For privacy scores, this means that the computation of the score should not depend only on the matrix R coming from a single social network. Several extensions of the original privacy score metric are possible based on the background knowledge type and source. In Sect. 4 we propose a new method to compute privacy scores, one that considers information about users beyond the ones in the social network, namely from a second/other social networks or more generally from the web.

Finally, it needs to be mentioned that the proposed privacy score metric measures only some aspect of privacy, namely attribute (item) disclosure and identity disclosure

arising from the attribute disclosure. There are several other aspects that may be of concern to the users of a social network, such as:

- the risks of identity disclosure that is not based on attribute disclosure—for example, based on behavioral observations,
- the risk of identity theft,
- the risk of link or relationship disclosure,
- the risk of group membership disclosure, or
- the risk of digital stalking.

How to measure these risks and help the users making informed decisions by presenting them a score reflecting these risks is still an open problem.

3.3 A Discussion Forum

The computation of privacy scores proposed by Liu and Terzi [4, 5] introduced in Sect. 2.1 assumes the analyzed information to be readily available for inclusion in the matrix R. As we noted in Sect. 3.2, non-structured information cannot be always easily included for analysis. It may be either information that is hard to extract—for example, relationships from photos—or previously not defined information—for example, non-structured text in natural language may contain multiple private items some of which may not be pre-defined as items of the matrix R.

Together with my Master's student Ján Žbirka we performed a few experiments [12], where simple natural language analysis was used to determine if some private information has been included in discussion comments on a news website. Since the users usually post multiple comments, they may contain multiple private information that must be looked-up for inclusion in the privacy scores. In our experiments, shown in the next section, we concentrated on information about political orientation before election and religious believes.

3.3.1 Experimental Results

Discussions of the Slovak news web site www.sme.sk were analyzed just before the government election in March 2012. From all the users that posted comments on the website, 5,268 users who posted more than 500 comments over the lifetime of the website were considered as the most active ones. In the three weeks before the election, these 5,268 most-active users posted 43,035 comments that were analyzed. Almost 20 % of the analyzed users revealed in their comments which political party in particular they were or were not going to vote for.

Summary of the findings are in Table 1 and all the other details about the experiment can be found in the Master's thesis of Ján Žbirka [12].

Since discussions on this website about religion and church are very heated, we also analyzed whether it is possible to find out the faith/religious beliefs of the users from their comments. The experiment that was done on the same sample of the users and comments have shown that simple natural language analysis can determine

Table 1 The number of the users (from the total of 5,268) who were find to disclose this information in discussion comments

Users who will	Vote	Not vote
At all	763 (14.5 %)	173 (3.3 %)
For a right wing party	209 (4.0 %)	194 (3.7 %)
For a left wing party	59 (1.1 %)	46 (0.9 %)
For a particular party	688 (13.1 %)	335 (6.4 %)

faith, although the users were more conservative in revealing their religious believes compared to the political orientation. In total, 133 (2.5 %) users were found to disclose their religion, and 106 (2.0 %) users were found to disclose that they are atheists.

4 Privacy Score Extension

The biggest disadvantage of the privacy scores that were outlined in Sect. 2.1 is the non-consideration of background knowledge. *Background knowledge* (sometimes referred to as external knowledge or auxiliary information) is some information about an individual that by itself is not a privacy disclosure, but combined with other information it becomes one. We propose two possible extensions of the original privacy score metric that take public background knowledge into account.

It needs to be noted that the reason to include background knowledge in the computation of the privacy score is two-fold. On the one hand, such extended privacy score will more precisely present users with privacy risks arising from publishing their information. On the other hand, using background knowledge also reduces another shortcoming of the original privacy scores. Namely, the more background knowledge is considered, the closer is the sensitivity of items to the true sensitivity. In other words, adding background knowledge to the privacy score computation also reduces or eliminates sensitivity paradoxes—see Sect. 3.2.

Our extended privacy score metric uses the same formula as in the Eq. (1) with sensitivity and visibility as the original privacy scores, but the information that is used to compute these—the matrix R—is extended by additional knowledge. We discuss two instances of this extension. The first one, presented next, combines information from two or several social networks when evaluating privacy risk of a user. The second instantiation of the extended privacy score metric, which we present in Sect. 4.2, uses "all the human knowledge" in privacy risk evaluation.

Our proposal of a simple inclusion of additional information in the privacy score computation is based on users' information (items) from multiple social networks. Let N be the number of considered social networks, and let R_t be as the already defined matrix R for a social network t, with $t = 1, \ldots, N$. Hence, R_t is a $n \times m$ matrix, where $R_t(i, j)$ represents the publicity of an item i for a user j—that is, non-disclosure when $R_t(i, j) = 0$ or disclosure when $R_t(i, j) > 0$ and possibly how far from the user j is the item public in the (graph of the) social network t.

It is likely in practice that not all the users are in every social network and that every item is in each of the corresponding profiles. Here we assume that the range of the items $i = 1, \ldots, n$ and the range of the users $j = 1, \ldots, m$ are the supersets over all the social networks, and so $R_t(i, j) = 0$ if an item i or user j do not exist in the social network t. We define the matrix R used for sensitivity and visibility computation as $R(i, j) = \max_t R_t(i, j)$ and use the formula from the Eq. (1) to compute the privacy score. Such privacy score better estimates the risk of privacy disclosure.

Together with my Master's student Lucia Maringová we performed a few experiments [13], where the same users on two social networks were evaluated for their privacy risks. The two social networks were of different type, so it was expected that the users would behave differently and therefore would disclose different amount of information about themselves on each social network. In our experiments, shown in the next section, we focused on computing privacy scores from each social network individually and then comparing the behavior of people in the terms of private information disclosure on two social networks.

4.1 Experimental Results

The purpose of the experiment is to show that privacy risks, as measured by the original and extended privacy score, are higher when two social networks are combined. Specifically, this means that some users tend to be conservative in one social network while publicly disclose private information in another social network.

For the experiments, profiles from the same users on two social networks were downloaded and analyzed. The social networks (websites) were Pokec.sk and Zoznamka.sk. They both belong to the same content provider/operator, and so use the same user authentication, which facilitated the pairing of the users from the two networks. Zoznamka.sk is a dating website, where a profile can contain up to 5 items: age, body type, weight, height, and contact. Pokec.sk is a website about chatting, messaging, and picture sharing. A profile on Pokec.sk can contain up to 34 items.

A sample of 3,923 users was selected. From all these users, there are only 23 users ($<1\,\%$) who completely filled all profile items on both websites. These people probably do not understand the risks of disclosing private information or ignore these risks, whether consciously or unconsciously by making a mistake. Roughly 32% of the users shared the same information on both sites.

Because of the nature of the website, users on the dating website Zoznamka.sk revealed more personal information about themselves. This is likely due to the fact that the users tried to create interest and attract the users who viewed their profiles. No user had less than 2 items (out of 5) filled on Zoznamka.sk. Conversely, many users on Pokec.sk left their profiles empty. What is of interest to us are the users who had empty profiles on Pokec.sk and non-empty profiles on Zoznamka.sk. Table 2 summarizes these users. All the details can be found in the Master's thesis of Lucia Maringová [13].

Table 2 The number of users who shared nothing on Pokec.sk, but had non-empty profiles on Zoznamka.sk Note that minimum items filled in on Zoznamka.sk was 2

On Pokec.sk	On Zoznamka.sk	# of users
0 items	2 items	542
0 items	3 items	107
0 items	4 items	961
0 items	5 items	119
0 items	>0 items	1,727
		44 %

In the terms of the privacy score, the users on Pokec.sk who had empty profiles would receive the score of 0, because they do not share or disclose anything. However, this would be awfully wrong in any privacy risk analysis, because private information about these users is publicly available and linkable to these users. At least two additional items can be learned about roughly 44 % of the users with empty profiles on Pokec.sk when considering Zoznamka.sk, so the extended privacy score computed over both networks for these users will be non-zero. This simple experiment itself shows the need to extend the original privacy scores from analyzing information over one social network to analyzing also auxiliary information.

4.2 Using All the Human Knowledge in Privacy Score Computation

Extending the original Privacy Scores by Liu and Terzi [4, 5] to multiple social networks certainly helps in privacy risk evaluation. The selection of social networks included in the extended privacy score computation, presented above, strongly impacts the quality and truthfulness of the score. The most truthful privacy risk evaluation can be achieved if all the human knowledge is used in the computation of the privacy score.

Including "all the human knowledge" in any computation is obviously impossible, so an approximation would have to suffice for all practical purposes. To effectively include the knowledge, we need to be able to quickly search for particular information or relation. Thus, we should use all the *indexed* human knowledge. Private databases and the "deep web" are believed to contain much more information than what is publicly available. In general, private information is out of reach for privacy adversaries as well as for privacy evaluators. Hence, we foresee to use all the *indexed public* human knowledge in the privacy score computation. Currently, the best instance and the best source of all the indexed public human knowledge is (Google) web search. In fact, there exists a proposal, namely Web-Based Inference Detection [14, 15], which takes advantage of the assumption that the web search is the proxy for all human knowledge.

Our idea is as follows: If an item of a user is not disclosed in the social network, we want to determine if the item has been disclosed elsewhere by using an inference

detection based on the other disclosed items for the user. Our inference detection method is heavily influenced by the Web-Based Inference Detection [14]. So, our idea rewritten in the terms of inference detection is: If there is a privacy-impacting inference detected for an undisclosed item, then this detected inference should be included in the privacy score computation.

More formally, we propose the following method to compute the privacy score:

Consider a social network of m users each having a possibility to fill a profile of n items. Let R be, as before, the $n \times m$ matrix over $\{0, 1\}$ with $R(i, j)$ representing whether the user j has (or has not) disclosed the item i. Let P be $n \times m$ array of strings with $P(i, j)$ being the value of the item i for the user j, in case this value has been disclosed. Let the set D_i be the domain of the item i. Finally, let β, γ, and δ be positive integers, where β and γ are parameters of the proposed algorithm that control the search depth, and δ is the parameter that controls the number of the most frequent words to be considered. Then the algorithm to extend R and determine the users' disclosures outside the social network is as follows:

For each user j, $j \in \{1, \ldots, m\}$

1. Let $S_j = \{k \mid R(k, j) = 1,\ k = 1, \ldots, n\}$ be the set of all disclosed items for the user j.
2. For each undisclosed item i, that is, for all $i \in \{1, \ldots, n\}$ with $R(i, j) = 0$
   (a) Let T be an empty multiset.
   (b) Take every subset $S'_j \subseteq S_j$ of size $|S'_j| \leq \beta$.
   (c) For every such subset $S'_j = \{i_1, \ldots, i_\ell\}$ with $\ell \leq \beta$
      (i) Use a web search engine to search for keywords $P(i_1, j), \ldots, P(i_\ell, j)$
      (ii) Retrieve the top γ most relevant documents containing these keywords
      (iii) Extract the top δ most frequent words from all these γ documents
      (iv) Add the top δ most frequent words to T together with their frequencies
   (d) Take the most frequent word from T that is also in D_i, if it exists.
   (e) If there is such word, let $R(i, j) = 1$.

After this, the newly enhanced matrix R contains the users' disclosures not just from the social network itself, but also from the web. Visibility and sensitivity values can be then computed from this matrix R, and the privacy score can be computed for each user using the Eq. (1).

The parameters β, γ, and δ can be tuned to achieve different trade-offs between the running time of the algorithm and the completeness and quality of the disclosure detection. In fact, these values can be different for different users, perhaps based on the number of items disclosed in the social network. Additional tuning can be achieved by performing the steps of the algorithm only for those users that have disclosed a "sufficient" number of items in the profile that would allow the web search to identify additional items.

5 Related Work

Our work was influenced by the approach by Liu and Terzi [4, 5], which provides users with a quantification of privacy risks due to sharing their profiles in a social network. Each user is assigned a privacy score based on their and all other users' profile items. The proposal is for a single social-network site, that is, a closed system evaluation of privacy that lacks the consideration and inclusion of background knowledge in computation of the privacy scores. We overcome this shortcoming by including background knowledge in the computation of privacy scores, see Sects. 3.3 and 4.

Privacy Scores is just one metric to help users understand their privacy risks. In Sect. 2, we have surveyed few other measures, scores, and tools—namely, Privacy Quotient and Privacy Armor [7], Privacy-Functionality Score [2], Privacy Index PIDX [9], PrivAware [6], Privometer [10], Privacy Wizard [1], and PViz [11].

The privacy risks of social-network sites are summarized in [16] and more recently in a survey [3]. Several papers present (relationship) privacy attacks in social networks [6, 17–21] or try to lower privacy risks and prevent privacy attacks in social networks [22, 23]. In addition, there are privacy risks from being tracked while browsing these websites [24].

Some form of background knowledge is usually considered in privacy attacks and is very likely available to attackers. Absolute privacy is impossible, because there will be always some background knowledge [25]. Inference techniques can then be used to attack or to help protect private data. In particular, web-based inference detection [14, 15] has been used to redact documents and prevent privacy leaks.

6 Conclusions

As more and more users are joining and using social-network web sites, they become more heavily used and their owners look for new ways to share different content, including private information and information that may lead to unwanted privacy leakages. It becomes increasingly difficult for individuals to control and manage their privacy in the vast amount of information available and collected about them.

Metrics, scores, and tools were proposed to facilitate social network users with a view of their privacy risks. In particular, Privacy Scores is a metric that presents users with a score that reflects their privacy risks arising from disclosing information in their profiles on a social network. We presented several shortcomings of the privacy scores as research opportunities for extending the privacy score metric. Next, we supported the need for extensions by experimental results from different websites and social networks. Finally, we proposed two extensions of the privacy score metric that consider additional background information about the users in the computation of the scores. Our approach provides a better decision support for individuals than the original privacy scores. Based on our extended privacy score metric, the users can compare their privacy risks with other fellow individuals and make informed

decisions about whether they share too much potentially private and sensitive information.

Acknowledgments This work started while the author was with the Universitat Rovira i Virgili, Catalonia and was partly funded by the Spanish Government through projects TSI2007- 65406-C03-01 "E-AEGI" and CONSOLIDER INGENIO 2010 CSD2007-00004 "ARES", and by the Government of Catalonia through grant 2009 SGR 1135. This work was also partly funded by the Slovak grant VEGA 1/0173/13 while the author was with the Slovak University of Technology. Final thanks go to my former Master's students Lucia Maringová and Ján Žbirka for carrying out the experiments.

References

1. Fang, L., LeFevre, K.: Privacy wizards for social networking sites. In: Proceedings of the 19th International Conference on World Wide Web, WWW 2010, pp. 351–360. ACM (2010)
2. Domingo-Ferrer, J.: Rational privacy disclosure in social networks. In: Proceedings of the 7th International Conference on Modeling Decisions for Artificial Intelligence, MDAI 2010, pp. 255–265. Springer (2010)
3. Zheleva, E., Getoor, L.: Privacy in Social Networks: A Survey. In: Social Network Data Analytics, pp. 277–306. Springer (2011)
4. Liu, K., Terzi, E.: A Framework for computing the privacy scores of users in online social networks. In: Proceedings of the Ninth IEEE International Conference on Data Mining, ICDM 2009, IEEE pp. 288–297. (2009)
5. Liu, K., Terzi, E.: A framework for computing the privacy scores of users in online social networks. ACM Trans. Knowl. Discov. Data **5**(1) (2010). (article no.6.)
6. Becker, J., Chen, H.: Measuring privacy risk in online social networks. In: Web 2.0 Security & Privacy 2009 Workshop of 2009 IEEE Symposium on Security and Privacy, W2SP 2009, IEEE (2009)
7. Srivastava, A., Geethakumari, G.: Measuring privacy leaks in online social networks. In: Proceedings of the 2013 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2013, pp. 2095–2100. (2013)
8. Sramka, M.: Privacy scores: assessing privacy risks beyond social networks. Infocommunications J. **4**(4), 36–41 (2012)
9. Nepali, R.K., Wang, Y.: Sonet: A social network model for privacy monitoring and ranking. In: Proceedings of the 2013 IEEE 33rd International Conference on Distributed Computing Systems Workshops, IEEE, pp. 162–166. (2013)
10. Talukder, N., Ouzzani, M., Elmagarmid, A.K., Elmeleegy, H., Yakout, M.: Privometer: privacy protection in social networks. In: Proceedings of the 2010 IEEE 26th International Conference on Data Engineering Workshops, ICDEW 2010, IEEE, pp. 266–269. (2010)
11. Mazzia, A., LeFevre, K., Adar, E.: Pviz comprehension tool for social network privacy settings. In: Proceedings of the 8th Symposium on Usable Privacy and Security. ACM (2012) (article no. 13.)
12. Zbirka, J.: Privacy risks arising from publishing private information on the web (in Slovak). Master's thesis, Advisor: Michal Sramka, Slovak University of Technology (2012)
13. Maringova, L.: Privacy risks arising from publishing private information in social networks (in Slovak). Master's thesis, Advisor: Michal Sramka, Slovak University of Technology (2012)
14. Staddon, J., Golle, P., Zimny, B.: Web-based inference detection. In: Proceedings of the 2007 USENIX Annual Technical Conference, USENIX 2007, USENIX Association, pp. 71–86. (2007)

15. Chow, R., Golle, P., Staddon, J.: Inference detection technology for Web 2.0. In: Web 2.0 Security & Privacy 2007 Workshop of 2007 IEEE Symposium on Security and Privacy, W2SP 2007, IEEE (2007)
16. Rosenblum, D.S.: What anyone can know: the privacy risks of social networking sites. IEEE Secur. Priv. **5**(3), 40–49 (2007)
17. Narayanan, A., Shmatikov, V.: De-anonymizing social networks. In: Proceedings of the 30th IEEE Symposium on Security and Privacy, S&P 2009, IEEE, pp. 173–187. (2009)
18. Krishnamurthy, B., Wills, C.E.: On the leakage of personally identifiable information via online social networks. In: Proceedings of the 2nd ACM Workshop on Online Social Networks, WOSN 2009, ACM, pp. 7–12. (2009)
19. Zheleva, E., Getoor, L.: To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles. In: Proceedings of the 18th International Conference on World Wide Web, WWW 2009, ACM, pp. 531–540. (2009)
20. Korolova, A., Motwani, R., Nabar, S.U., Xu, Y.: Link privacy in social networks. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM 2008, ACM, pp. 289–298. (2008)
21. Backstrom, L., Dwork, C., Kleinberg, J.M.: Wherefore art thou R3579X?: anonymized social networks, hidden patterns, and structural steganography. In: Proceedings of the 16th International Conference on World Wide Web, WWW 2007, ACM, pp. 181–190. (2007)
22. Felt, A., Evans, D.: Privacy protection for social networking platforms. In: Web 2.0 Security & Privacy 2008 Workshop of 2008 IEEE Symposium on Security and Privacy, W2SP 2008, IEEE (2008)
23. Zheleva, E., Getoor, L.: Preserving the privacy of sensitive relationships in graph data. In: Proceedings of the First ACM SIGKDD International Workshop on Privacy, Security, and Trust in KDD, PinKDD 2007, ACM, pp. 153–171. (2007)
24. McKinley, K.: Cleaning up after cookies. Technical report, iSEC Partners (2008)
25. Dwork, C.: Differential privacy. In: Proceedings of the 33rd International Colloquium on Automata, Languages and Programming, ICALP 2006, pp. 1–12. Springer (2006)