Privacy Risk and Data Utility Assessment on Network Data

Roberto Pellungrini(✉)

Department of Computer Science, University of Pisa, Pisa, Italy
roberto.pellungrini@di.unipi.it

Abstract. In the modern Internet era the usage of social networks such as Twitter, Instagram and Facebook is constantly increasing. The analysis of this type of data can help us understand interesting social phenomena, because these networks intrinsically capture the new nature of user interactions. Unfortunately, social network data may reveal personal and sensitive information about users, leading to privacy violations. In this paper, we propose a study of privacy risk for social network data. In particular, we empirically analyze a set of privacy attacks on social network data by using the privacy risk assessment framework PRUDEnce. After simulating the attacks on real data, we first analyze how the privacy risk is distributed over the whole population. Then, we study the effect of high-risk users sanitization on some common network metrics.

Keywords: Privacy · Attack models · Social networks

1 Introduction

Social networks are used by people everyday for different purposes: for interacting with friends (Facebook), for professional activities (LinkedIn), for spreading information, news and multimedia material (Twitter and Instagram). Nowadays, the analysis of social network data is fundamental to study and understand social phenomena. The social network analysis can help in understanding customer interactions and reactions [15], marketing strategies based on communities or singles users, migration flows, fake news diffusion or virus spread [16], etc. However social network data may contain sensitive and private information about the real people that actively operate in the network. Therefore, different techniques have been used to anonymize the data, the simplest way being replacing identity with pseudonymous keys. However, Backstrom et al. [3] showed that this technique is not enough for privacy protection as malicious adversaries still may succeed in re-identifying individuals using a background knowledge attack.

In order to enable a practical application of the privacy-preserving techniques proposed in the literature, Pratesi et al. in [14] proposed PRUDEnce, a framework for systematic privacy risk assessment. This framework follows the idea of the EU General Data Protection Regulation, which explicitly imposes on data controllers an assessment of the impact of data protection for the most risky

J. Bowles et al. (Eds.): DataMod 2021, LNCS 13268, pp. 93–106, 2022.
https://doi.org/10.1007/978-3-031-16011-0_7

processes.[1] In [14], Pratesi et al. show the applicability of their framework on mobility data. In this paper, we propose to apply PRUDEnce framework for the privacy risk assessment in social network data. This requires to first formally define a set of privacy attacks on social network data, then simulate them on real data to empirically evaluate the individual privacy risks, and then, evaluate the data utility by considering only non-risky data. In order to evaluate the data utility, we perform an analysis that highlights the degradation of the social network structure in case we only consider non-risky nodes and their connections.

The paper is organized as follows. In Sect. 2, we discuss some of the related works in the literature. In Sect. 3, we define the data structures to describe social network data according to different data aggregations. In Sect. 4, we introduce the framework used for the privacy risk assessment. In Sect. 5, we formally define the privacy attacks on social network data. In Sect. 6, we show the results of our experiments on the attack simulations. In Sect. 8, we present an analysis on the network degradation, discusses some related work. In Sect. 9, we draw our conclusions and discuss future works.

2 Related Work

The concept of privacy-by-design was initially developed by Ann Cavoukian [5] to address the ever-growing and systemic effects of Information and Communication Technologies, and of large-scale networked data systems in the 90's. This concept basically expresses the general approach of embedding privacy into the design, development and management of information. A related study on the application of the concept of privacy-by-design to social media is [7] where the authors develop a social networking privacy framework and privacy model for applying privacy-by-design principles to social networks for both desktop and mobile devices. This approach mitigates many current social network privacy issues, and leads to a more controlled form of privacy assessment.

One of the classical works in the field of privacy risk assessment is the LINDDUN methodology [6], a privacy-aware threat analysis framework based on Microsoft's STRIDE methodology [19]. PRUDEnce [14] builds on these principles, developing a privacy risk assessment framework applicable to any kind of data [12]. While the models and methodology presented in these works have been used previously on human mobility data, they are flexible enough to be adapted to social network data. For modeling the attacks, we rely on the contributions made in [1,9,17,20,21]. We apply the general structure of these attacks, tweaking some of them to our specific need, as explained in Sect. 5.

Privacy for social media networks is a high interest topic, as show in works such as [11] where the authors highlight how privacy awareness changes the perspectives and motivations of users of a social media. In the context of privacy for online social networks Liu and Terzi [10] propose a framework for computing privacy scores for each user in the network. Such scores indicate the potential risk caused by her participation in the network. Our effort in defining possible

[1] The EU General Data Protection Regulation can be found at http://bit.ly/1TlgbjI.

attacks and studying their applications on real network goes in the direction of offering more tools to actually provide realistic evaluation of privacy risk to individuals. In [4] Becker and Chen propose a framework called PrivAware, a tool to detect and report unintended information loss in online social networks. In [2] Ananthula et al. discuss a "Privacy Index" (PIDX) used to measure a user's privacy exposure in a social network. They have also described and calculated the "Privacy Quotient" (PQ) i.e. a metric to measure the privacy of the user's profile using a naive approach. Pensa and Blasi in [13] have proposed a supervised learning approach to calculate a privacy score of an individual in social network data based on the actual people allowed to access the profile of the individual.

3 Data Definitions

Social network have traditionally been modeled as graphs:

Definition 1 (Social Network). *We model a social network as a simple graph* $G = (V, E, L, \Gamma)$*, where* V *is the set of vertices representing individuals,* $E \subseteq V \times V$ *is the set of edges representing the relationships between individuals,* L *is a set of labels, and* $\Gamma : V \rightarrow l$ *is a labeling function that maps each vertex to a subset of labels* l *with* $l \subseteq L$*.*

To keep our definition simple, we assume that edges do not have any labels. In a social network, the direction of an edge indicates the relationship between vertices and can be used to distinguish the type of relationship: single-sided or mutual. For our purposes, we will assume that all relationships are mutual. From the social network graph we can derive data structures representing aggregated information. These are used to expose less information while still enabling the computation of standard network metrics. Clearly, this data transformation helps privacy preserving analyses and the respect of data minimization principle.

Definition 2 (Friendship Vector). *The friendship vector* F_v *of an individual* $v \in V$ *is a set of vertices* $F_v = \langle v_1, v_2 \ldots, v_n \rangle$ *representing individuals connected to* v *in the social network graph.*

The friendship vector of a node v essentially represents the neighborhood of the individual v at distance 1.

Definition 3 (Label vector). *The label vector of an individual* v *is a set of labels* $LA_v = \langle la_1, la_2 \ldots, la_m \rangle$*. Each* $la_j = (f, l)$ *(with* $j \in \{1, 2, \ldots, |L|\}$*) is a pair composed of a feature name* f *and the associated label value* l*. The label vector of an individual can be empty.*

Each label describes a profile feature of an individual, for example *gender*: *'female'* or *'male'*, *educational information: 'Pisa University'* or *'Stanford University'*, etc.

Definition 4 (Degree vector). *The degree vector of an individual v, denoted by $D_v = \langle d_{v_1}, d_{v_2}, \ldots, d_{v_n} \rangle$, represents the number of friends of each friend of v. Thus, each element d_{v_i} is equal to the length of the friendship vector of the individual v_i in the social network graph, i.e., $d_{v_i} = |(F_{v_i})|$.*

Definition 5 (Mutual Friendship vector). *The mutual friendship vector of an individual v, denoted by $MF_v = \langle mf_1, \ldots, mf_n \rangle$, represents the number of common friends of v with each one of its friends v_i. Thus, each element mf_i is equal to the cardinality of the intersection between the friendship vector of v and the one of v_i, i.e., $mf_i = | F_v \cap F_{v_i} |$.*

Taking in consideration all of the structures defined above we can define a Social Network Dataset as follows:

Definition 6 (Social Network Dataset). *A social network dataset is a set of data structures $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_k\}$ where $\mathcal{S}_v$ $(1 \leq v \leq k)$ is the social network data structure of an individual v.*

A Social Network Dataset represents a possible aggregation of a social network that some data provider may share or publish for some usage. A malicious adversary can attack a Social Network Dataset using some previously acquired knowledge about one or more individuals in the dataset, i.e., a background knowledge.

4 Privacy Risk Assessment Framework

Given the rapid growth in the number of services and applications based on social networks, there is increasing concern about privacy issues in published social network data. The prevention of node/individual re-identification is one of the critical issues. With some background knowledge about an individual in a social network, an adversary may perform a re-identification attack and disclose the identity of the individual. To preserve privacy, it is not sufficient to remove all identifiers, as shown in [20,21]. In this paper we want to empirically study the privacy risk in social network data using the framework proposed in [14]. PRUDEnce is a system enabling a privacy-aware ecosystem for sharing personal data. The main components of its architecture are shown in Fig. 1. The typical scenario considered is one where a Service Developer (SD) requests personal data, such as social network data, from a Data Provider (DP) to develop services or perform an analysis. The Data Provider has to guarantee the right to privacy of the individuals whose data are recorded. Thus, the data stored by DP cannot be shared directly without assessing the privacy risk of the individuals represented in the data. Once privacy risk has been assessed, DP can choose how to protect the data before sharing, selecting the privacy preserving methodology most appropriate for the data to be shared. Assuming that the Data Provider stores a database D, it aggregates, selects, and filters the dataset D to meet the requirements of the Data Analyst and produces a set of social network datasets

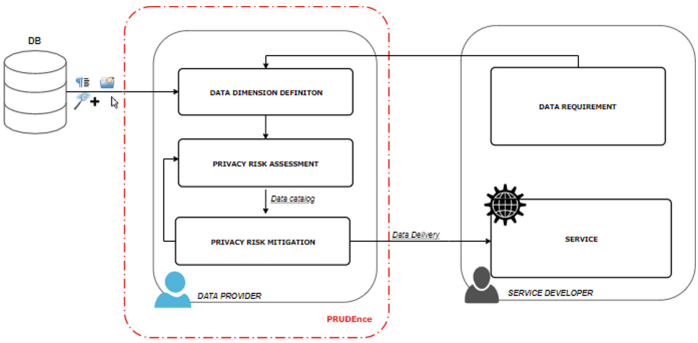


Fig. 1. PRUDence: privacy-aware data sharing ecosystem

$\{S_1, S_2, \ldots, S_k\}$ each with a different data structure and/or aggregation of the data. The Data Provider then performs the privacy risk assessment.

The privacy risk assessment component of the framework has to produce a quantitative measure of privacy risk. Such measure depends on the kind of attack simulated, the kind of data, and on the aggregation on the data itself. The simulation of a privacy attack takes place in two phases: first, we assume that a malicious adversary gathers, in some way, a *background knowledge* about an individual (e.g., a part of their friendship vector) and then the adversary uses the acquired background knowledge to re-identify the individual in the social network dataset. Every background knowledge can be configured in many ways, and for each configuration there can be many background knowledge instances. To explain how risk is computed, we give the formal definitions of these concepts:

Definition 7 (Background knowledge Category). *A background knowledge category B of an attack is the type of information known by the malicious adversary. It represents the dimensions of data considered by the adversary, i.e., the knowledge of the friendship vector, or the neighboring vector etc.*

Definition 8 (Background Knowledge Configuration). *A background knowledge configuration $B_k \in \mathcal{B} = \{B_1, B_2, ..., B_n\}$ represents the k elements of the background knowledge category $\mathcal{B}$ known to the adversary. For example, the adversary might know $k = 3$ of the friends in the friendship vector of an individual.*

Definition 9 (Background Knowledge Instance). *A background knowledge instance $b \in B_k$ is a specific information known by the adversary, i.e., the actual portion of data structure known by the adversary.*

As an example, suppose that an adversary has, as background knowledge category, the friendship vector of a user, and suppose $F_v = \langle v_1, v_2, v_3, v_4 \rangle$. If the background knowledge configuration that we assume for the adversary is B_2, a possible instance could be $b = v_1, v_4$ or $b = v_3, v_4$ for example. Each instance

gives to the adversary a probability of re-identifying the individual v in the dataset.

Definition 10 (Probability of Re-identification). *Given an attack, a function matching(s, b) indicating whether or not a record $s \in S$ matches the instance of background knowledge configuration $b \in B_k$, and a matching set $M(S, b) = \{s \in S \mid matching(s, b) = True\}$, we define the probability of re-identification of an individual v in dataset S as:*

$$PR_S(s = v|b) = \frac{1}{|M(S, b)|}$$

that is the probability of correctly linking the data structure $s \in S$ to v given the background knowledge instance b.

Note that $PR_S(s = v|b) = 0$, in case an individual v does not belong to S.

PRUDEnce is a worst-case scenario framework, so when simulating an attack we have to assume that the adversary has access to the worst possible background knowledge instance. We take the maximum probability of re-identification among all $b \in B_k$ as risk of the re-identification risk for that individual:

Definition 11 (Risk of re-identification). *The risk of re-identification of an individual v is $Risk(v, S) = \max PR_S(s = v \mid b), \forall b \in B_k$, i.e., the maximum probability of re-identification.*

Our definition of probability of re-identification and privacy risk derives from the work of Sweeney in [18].

To better understand these concepts, we provide an example of risk computation definitions of probability and risk of re-identification.

Let us consider a set of individuals (nodes) $V = \{1, 2, 3, 4, 5, 6, 7\}$ and the corresponding dataset S composed of the friendship vectors of individuals:

$F_1 = \langle$'3', '4', '6'$\rangle$ $F_2 = \langle$'4', '6'$\rangle$
$F_3 = \langle$'1', '5', '7'$\rangle$ $F_4 = \langle$'1', '2', '6'$\rangle$
$F_5 = \langle$'3', '7'$\rangle$ $F_6 = \langle$'1', '2', '4', '7'$\rangle$
$F_7 = \langle$'3', '5', '6'$\rangle$

Let us assume an adversary wants to perform an attack on individual 6 and knows two friends of that individual. The background knowledge configuration in this case is B_2. We compute the privacy risk of the individual 6, given the dataset S of friendship vectors and the knowledge of the adversary as follows:

1. We compute every possible instance $b \in B_2$ which are: {(1, 2), (1, 4), (1, 7), (2, 4), (2, 7), (4, 7)}
2. We compute the probability of re-identification for each background knowledge instance, matching it with the dataset S and counting the matching individuals. For example, the first instance $b = (1, 2)$ has the probability of re-identification $PR_S(s = 6|(1, 2)) = \frac{1}{2}$ because both 4 and 6 include $b = (1, 2)$ in their friendship vectors. We do this for every instance, obtaining the following values: $\frac{1}{2}, 1, \frac{1}{2}, 1, 1, 1$.
3. We take the maximum probability of re-identification as risk for individual 6: $Risk(6, S) = max(\frac{1}{2}, 1, \frac{1}{2}, 1, 1, 1) = 1$

5 Privacy Attack on Social Networks

Given the privacy framework we presented, the definition of an attack depends entirely on the matching function used to understand if a particular background knowledge instance can be found in the data structure of an individual. In this section, we describe the different type of attacks detailing their matching function.

5.1 Neighborhood Attack

The *neighborhood attack* considers an adversary who only knows a certain number of friends/neighbors of an individual. More technically, the adversary has an information about the nodes which are connected to the victim node in the social network graph. This type of attack was introduced in [20]. Background knowledge instances for this kind of attack are portions of the friendship vector F_v of an individual.

Definition 12 (Neighborhood Attack Matching). *Given the instance b, we define the matching function of the neighborhood attack as follows:*

$$Matching(b, F_v) = \begin{cases} true & b \subseteq F_v \\ false & otherwise \end{cases} \tag{1}$$

5.2 Label Pair Attack

The *label pair attack* considers an adversary who knows a certain number of pairs of features with their values of an individual. Each label pair in key-value format $la_i = (f, l)$ is distinct in a label vector of an individual. Similar type attack has been defined in [9] by using the label pair knowledge on two connected nodes. In our work, we consider that an adversary uses label pair knowledge of just one individual and it may be sufficient for his re-identification within S. Therefore, background knowledge instances for this kind of attack are portions of the label pair vector LA_v of an individual.

Definition 13 (Label Pair Attack Matching). *Given the instance b, we define the matching function of the label pair attack as:*

$$Matching(b, LA_v) = \begin{cases} true & b \subseteq LA_v \\ false & otherwise \end{cases} \tag{2}$$

5.3 Neighborhood and Label Pair Attack

Starting from the previous two attacks, we define a new and stronger attack that we call *neighborhood and label pair attack*. In this case, we consider an adversary knowing a certain number of friends/neighbors and a certain number of feature labels of an individual at the same time. In other words, it combines the background knowledge of the two previous attacks.

Definition 14 (Neighborhood and Label Pair Attack Matching). *Given the instance $b = (b', b'')$, we define the matching function of the neighborhood and label pair attack as:*

$$Matching(b, F_v, LA_v) = \begin{cases} true & b' \subseteq F_v \ \wedge \ b'' \subseteq LA_v \\ false & otherwise \end{cases} \tag{3}$$

5.4 Friendship Degree Attack

In a *friendship degree attack*, the adversary knows the degree of a number of friends of the victim as well as the degree of the victim. This type of attack has been defined in [20]. A background knowledge instance for this kind of attack will be a portion of the degree vector D_v of an individual.

Definition 15 (Friendship Degree Attack Matching). *Given the instance b, we define the matching function of the friendship degree attack as:*

$$Matching(b, D_v) = \begin{cases} true & len(D_v) \wedge d \in D_v \forall d \in b \\ false & otherwise \end{cases} \tag{4}$$

5.5 Mutual Friend Attack

In a *mutual friend attack*, the adversary knows the number of mutual friends of the victim and some of its neighbors. This type of attack has been already defined in [17]. A background knowledge instance for this kind of attack will be a portion of the mutual friendship vector MF_v of an individual.

Definition 16 (Mutual Friend Attack Matching). *Given the instance b, we define the matching function of the mutual friend attack as:*

$$Matching(b, MF_v) = \begin{cases} true & b \subseteq MF_v \\ false & otherwise \end{cases} \tag{5}$$

5.6 Neighborhood Pair Attack

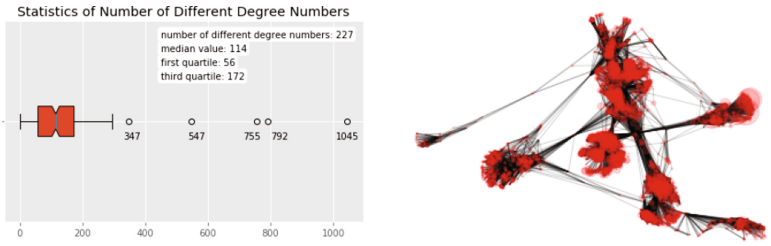
In a *neighborhood pair attack*, the adversary knows subset of the friends of the victim who are friends with each other, that is a subset of F_v in which v_i and v_j are connected to each other $v_i \in F_{v_j}$, $v_j \in F_{v_i}$ and $v_i, v_j \in F_v$. Similar type of attack has been already defined in [1]. With respect to the original definition, in our work, we reduce the knowledge of adversary by eliminating the degree knowledge of the victim about the individual, because we would like to simulate a less powerful kind of attack.

Definition 17 (Neighborhood Pair Attack Matching). *Given the instance b, we define the matching function of the neighborhood pair attack as:*

$$Matching(b, F_v) = \begin{cases} true & v_i \in F_{v_j} \wedge v_j \in F_{v_i} \wedge v_i, v_j \in F_v \forall (v_i, v_j) \in b \\ false & otherwise \end{cases} \tag{6}$$

6 Experimental Dataset

In our work, we use the Facebook Dataset provided by Stanford University's "Stanford Large Network Dataset Collection" [8]. This dataset includes node features (profiles), circles and ego networks. Nodes have been anonymized by replacing the Facebook-internal id's for each user with a new value. Feature vectors from this dataset have also been provided while the interpretation of those features has been anonymized. After aggregating all data, we obtain a social network graph of 4039 nodes and 88,234 edges. Roughly half of the all individuals have 30 friends/neighbors or less. In Fig. 2 we can see some visual information about the dataset.



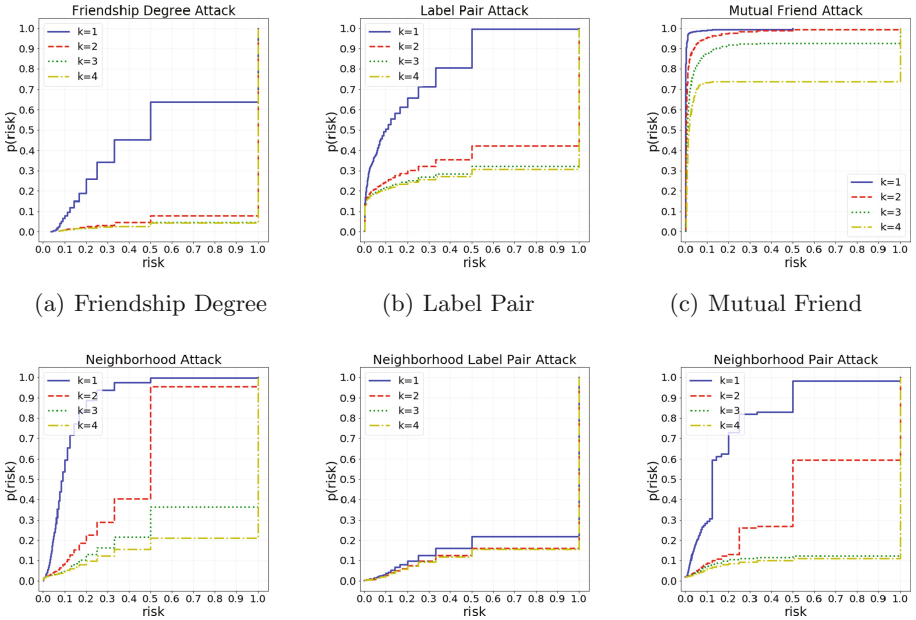
(a) Degree Statistics (b) Network Visualization

Fig. 2. Visual information about Stamford Facebook dataset

7 Privacy Risk Assessment Results

In this section, we present the results of the simulation of attacks defined in Sect. 5. We simulated all the defined attacks setting the background knowledge configuration value to $k = 1, 2, 3, 4$, that is with four different lengths of the adversary background knowledge. We discretized the privacy risk in six intervals: [0.0], (0.0, 0.1], (0.1, 0.2], (0.2, 0.3], (0.3, 0.5] and (0.5, 1.0] from the lowest privacy risk to the highest.

Figure 3 shows as privacy risk for the attacks on network data varies significantly. In general and as expected, the number of individuals in the highest risk level and lowest risk level increases while the background knowledge configuration value k increases. We can observe that, for most of the attacks, we reach a sort of plateau increasing k to values 3 and 4. This is a phenomenon observed also in other types of data [12]. The most interesting results can be seen for the neighborhood label pair attack: with respect to the simple label attack or neighborhood attack, the mixed attack leads to an increase of the number of high risk individuals by a great margin. The mutual friend attack is weaker with respect to all the others. Indeed, in each setup of the background knowledge configuration value k, many individuals belong to the privacy risk level $(0.0, 0.1]$. This is not surprising since the Mutual Friend attacks uses the number of mutual



(a) Friendship Degree (b) Label Pair (c) Mutual Friend

(d) Neighborhood (e) Neighborhood Label Pair (f) Neighborhood Pair

Fig. 3. Cumulative distributions of privacy risk for social network data.

friends of one node, which has a pretty even distribution over the entire network. Overall, the results suggest that basic topological information such as immediate neighborhood yield sufficient information for powerful privacy attacks, and that even a small amount of information can result in significant risk for the entire network.

8 Analysis on Network Degradation

Given the risk found in our assessment, we now remove from the data those individuals that are at or above a certain threshold of privacy risk and try to understand what impact will this have on the network: removing certain nodes from the network may lead to a disconnected networks. In these cases we compute the metrics and statistics on the biggest component of the unconnected network, thus we exclusively study the network's giant component. In order to verify the network degradation varying the minimum privacy risk guaranteed we compute: the total number of nodes preserved in the network, number of nodes preserved in the giant component, the total number of edges between individuals in the giant component and number of disconnected components generated.

We chose four thresholds of the privacy risk for all possible attacks. Thus, for each attack and background knowledge configuration value we created four distinct datasets which are:

- D_1: The dataset with all individuals
- $D_{0.5}$ The dataset with individuals whose privacy risk is between 0.0 and 0.5 included.
- $D_{0.33}$ The dataset with individuals whose privacy risk is between 0.0 and 0.333 included.
- $D_{0.25}$ The dataset with individuals whose privacy risk is between 0.0 and 0.25 included.

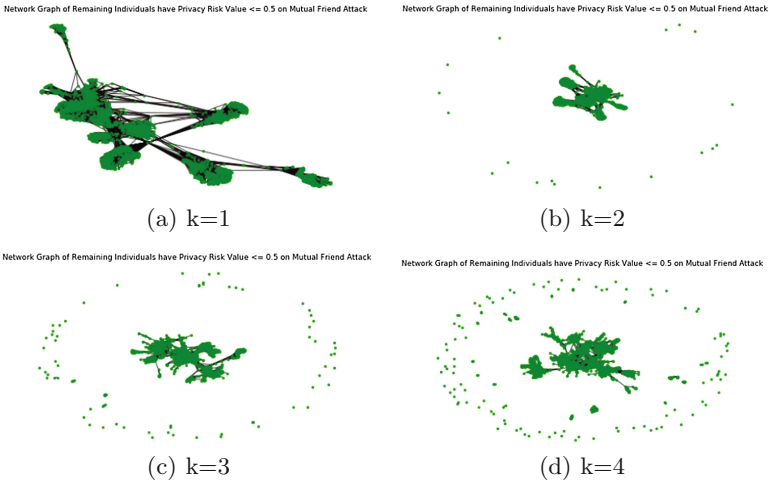
Tables 1 and 2 show the results of degradation analysis for the above statistics. For each attack the tables report the statistics for both $k = 1$ and $k = 4$ (the lowest and the greatest adversary knowledge). In these tables we indicate with **NA** the neighborhood attack, **LA** the label pair attack, **NLA** the neighborhood and label pair attack, **FDA** the friendship degree attack, **MFA** the mutual friend attack and **NPA** the neighborhood pair attack.

The results indicates that also considering an attack with a weak background knowledge the effect of taking into consideration only non-risky nodes has an important negative impact on the network quality. First of all, we can observe that for any attack there is an increment of the disconnected components in the networks (Table 1), that leads to a decrease in connectivity for the network. This effect is present even with weak attacks such as the *mutual friend attack*.

Tables 2 and 3 show that both the number of nodes in the whole network and in the giant component are mainly affected by the consequences of attacks based on the neighborhood information such as *neighborhood attack*, *friendship degree attack* and *neighborhood pair attack*.

Instead, Table 4 show that the number of edges is sensitive to the attacks that take into consideration the relationship between nodes such as the *friendship degree attack*.

Figure 4 shows what happens to the social network structure under the *mutual friend attack*, varying k. We can see how the network changes when we remove individuals with risk equal to 1 varying the background knowledge configuration. Even more evident effects can be seen for more powerful attacks.



(a) k=1 (b) k=2

(c) k=3 (d) k=4

Fig. 4. Cumulative distributions of privacy risk for social network data.

Table 1. Number of disconnected components varying privacy risk.

	NA		LA		NLA		FDA		MFA		NPA	
	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4
D_1	1	1	1	1	1	1	1	1	1	1	1	1
$D_{0.5}$	91	453	1	160	246	150	131	247	1	120	1	431
$D_{0.33}$	171	460	78	159	227	143	143	239	19	120	112	413
$D_{0.25}$	226	412	84	157	203	140	162	239	19	120	112	405

Table 2. Number of Nodes in the network varying privacy risk.

	NA		LA		NLA		FDA		MFA		NPA	
	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4
D_1	4039	4039	4039	4039	4039	4039	4039	4039	4039	4039	4039	4039
$D_{0.5}$	4029	907	4026	1282	883	976	2573	295	4039	2982	3971	511
$D_{0.33}$	3940	700	3251	1178	648	937	1836	276	4021	2982	3351	488
$D_{0.25}$	3792	579	2885	1131	498	925	1378	276	4021	2982	3313	472

Table 3. Number of nodes in the giant component varying privacy risk.

	NA		LA		NLA		FDA		MFA		NPA	
	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4
D_1	4039	4039	4039	4039	4039	4039	4039	4039	4039	4039	4039	4039
$D_{0.5}$	3732	25	4026	506	192	348	1426	7	4039	2226	3971	13
$D_{0.33}$	3507	23	2995	451	118	337	800	6	4003	2226	2566	13
$D_{0.25}$	3293	20	2618	429	83	335	326	6	4003	2226	2532	13

Table 4. Number of edges in the giant component varying privacy risk.

	NA		LA		NLA		FDA		MFA		NPA	
	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4	k = 1	k = 4
D_1	88234	88234	88234	88234	88234	88234	88234	88234	88234	88234	88234	88234
$D_{0.5}$	82305	88	87688	2940	545	1587	12310	21	88234	15160	87744	71
$D_{0.33}$	80335	223	46610	2506	851	1528	4527	15	82800	15160	30900	71
$D_{0.25}$	77368	163	34520	2276	211	1503	1753	15	82800	15160	30758	71

9 Conclusions and Future Works

Social network data are a precious proxy to improve our understanding of social dynamics. Nevertheless, it contains sensitive information which, if analyzed with malicious intent, can lead to privacy violations for the individuals involved. In this paper, we proposed to apply PRUDEnce framework for assessing privacy risk in social networks and for evaluating the network degradation in case we consider only non-risky individuals and their connections. Our study indicates that for social network, privacy attacks can yield high privacy risk, even when the background knowledge is based on aggregated structures. We also showed how basic sanitization of the network is difficult, due to several properties of the network being disrupted by the cancellation of high risk nodes. PRUDEnce demands high computational costs for attack simulations. In order to address this problem it would be interesting to investigate the possibility to apply machine learning methods able to learn the relationship between some graph properties and the privacy risk, in order to predict the node privacy exposure.

Acknowledgments. This work has been funded by the European projects SoBigData-PlusPlus (Grant Id 871042).

References

1. Abawajy, J.H., Ninggal, M.I.H., Herawan, T.: Vertex re-identification attack using neighbourhood-pair properties. Concurr. Comput. Pract. Exp. **28**(10), 2906–2919 (2016). https://doi.org/10.1002/cpe.3687
2. Ananthula, S., Abuzaghleh, O., Alla, N.B., Chaganti, S.P., Kaja, P.C., Mogilineedi, D.: Measuring privacy in online social networks. Int. J. Secur. Priv. Trust Manag. **4**(2), 01–09 (2015). https://doi.org/10.5121/ijsptm.2015.4201. www.airccse.org/journal/ijsptm/papers/4215ijsptm01.pdf
3. Backstrom, L., Dwork, C., Kleinberg, J.: 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, pp. 181–190. ACM, New York (2007). https://doi.org/10.1145/1242572.1242598
4. Becker, J., Chen, H.: Measuring Privacy Risk in Online Social Networks
5. Cavoukian, A.: Privacy by design the 7 foundational principles, August 2009. www.iab.org/wp-content/IAB-uploads/2011/03/fred_carter.pdf
6. Deng, M., Wuyts, K., Scandariato, R., Preneel, B., Joosen, W.: A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. Requir. Eng. **16**(1), 3–32 (2011). https://doi.org/10.1007/s00766-010-0115-7

7. Islam, M.B., Iannella, R.: Privacy by design: does it matter for social networks? In: Privacy and Identity Management for Life - 7th IFIP WG 9.2, 9.6/11.7, 11.4, 11.6/PrimeLife International Summer School, Trento, Italy, 5–9 September 2011, Revised Selected Papers, pp. 207–220 (2011)
8. Leskovec, J., Krevl, A.: SNAP Datasets: Stanford large network dataset collection, June 2014. http://snap.stanford.edu/data
9. Liu, C., Yin, D., Li, H., Wang, W., Yang, W.: Preserving privacy in social networks against label pair attacks. In: Ma, L., Khreishah, A., Zhang, Y., Yan, M. (eds.) WASA 2017. LNCS, vol. 10251, pp. 381–392. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60033-8_34
10. Liu, K., Terzi, E.: A framework for computing the privacy scores of users in online social networks. TKDD **5**(1), 6:1–6:30 (2010). https://doi.org/10.1145/1870096.1870102
11. Mvungi, B., Iwaihara, M.: Associations between privacy, risk awareness, and interactive motivations of social networking service users, and motivation prediction from observable features. Comput. Hum. Behav. **44**, 20–34 (2015). https://doi.org/10.1016/j.chb.2014.11.023
12. Pellungrini, R., Pappalardo, L., Pratesi, F., Monreale, A.: Analyzing privacy risk in human mobility data. In: Mazzara, M., Ober, I., Salaün, G. (eds.) STAF 2018. LNCS, vol. 11176, pp. 114–129. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-04771-9_10
13. Pensa, R.G., Di Blasi, G.: A semi-supervised approach to measuring user privacy in online social networks. In: Calders, T., Ceci, M., Malerba, D. (eds.) DS 2016. LNCS (LNAI), vol. 9956, pp. 392–407. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46307-0_25
14. Pratesi, F., Monreale, A., Trasarti, R., Giannotti, F., Pedreschi, D., Yanagihara, T.: PRUDEnce: a system for assessing privacy risk vs utility in data sharing ecosystems. Trans. Data Priv. **11**, 139–167 (2018)
15. Rossetti, G., Milli, L., Giannotti, F., Pedreschi, D.: Forecasting success via early adoptions analysis: a data-driven study. PLoS ONE **12**(12), e0189096 (2017)
16. Rossetti, G., Milli, L., Rinzivillo, S., Sîrbu, A., Pedreschi, D., Giannotti, F.: NDlib: a python library to model and analyze diffusion processes over complex networks. Int. J. Data Sci. Anal. **5**(1), 61–79 (2017). https://doi.org/10.1007/s41060-017-0086-6
17. Sun, C., Yu, P.S., Kong, X., Fu, Y.: Privacy preserving social network publication against mutual friend attacks. Trans. Data Priv. **7**(2), 71–97 (2014). www.tdp.cat/issues11/abs.a195a14.php
18. Sweeney, L.: k-anonymity: a model for protecting privacy. Int. J. Uncertainty Fuzziness Knowl.-Based Syst. **10**(05), 557–570 (2002). https://doi.org/10.1142/S0218488502001648
19. Swiderski, F., Snyder, W.: Threat Modeling. O'Reilly Media Inc., New York (2009). oCLC: 609857070
20. Tai, C., Yu, P.S., Yang, D., Chen, M.: Privacy-preserving social network publication against friendship attacks. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 21–24 August 2011, pp. 1262–1270 (2011)
21. Zhou, B., Pei, J.: Preserving privacy in social networks against neighborhood attacks. In: Proceedings of the 24th International Conference on Data Engineering, ICDE 2008, Cancún, Mexico, 7–12 April 2008, pp. 506–515 (2008)