AIUCD 2021
DH per la società: e-guaglianza, partecipazione, diritti e valori nell'era digitale
10° congresso annuale PISA 19-22 gennaio
DIGITAL PUBLIC HUMANITIES
OPEN CULTURE
RETI SOCIALI
TECH ECONOMY
E-PARTICIPATION
TECNOLOGIE ASSISTIVE

Versione PROVVISORIA del contributo presentato al Convegno Annuale

DISCLAIMER

Questa versione dell'abstract non è da considerarsi definitiva e viene pubblicata esclusivamente per facilitare la partecipazione del pubblico al convegno AIUCD 2021

Il Book of Abstract contenente le versioni definitive e dotato di ISBN sarà disponibile liberamente a partire dal 19 gennaio sul sito del convegno sotto licenza creative commons.

Privacy Risk Assessment on Network Data

Roberto Pellungrini[1], Anna Monreale[2]

[1] Department of Computer Science, University of Pisa, Italy – roberto.pellungrini(«»)di.unipi.it

[2] Department of Computer Science, University of Pisa, Italy – anna.monreale(«»)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. As a consequence, before applying any analytical framework on social network data, it is important to empirically assess the privacy risk of users to identify risky data to be treated with appropriate privacy-preserving technique. 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 analyze the distribution of privacy risk.

PAROLE CHIAVE

Privacy, Privacy Risk, Social Networks

1. INTRODUCTION

Social networks are used by people every day 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 [1], marketing strategies based on communities or singles users, migration flows, fake news diffusion or virus spread [2]. However social network data may contain sensitive and private information about the real people that actively operate in the network. Backstrom et al. [3] showed that basic anonymization techniques are 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. [4] 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 the responsibility of assessing privacy risk for data mining processes.[1] In [4] Pratesi et al. shows 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. We show how PRUDEnce can be applied to social network data and what kind of insights we can gather from the analysis of privacy risk in social networks. We show that even when a social network is represented with aggregative data structures, it is still possible to conduct background knowledge attacks and re-identify individuals in the network. 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.

2. DATA DEFINITION

Social networks have traditionally been modeled as graphs $G = (V, E, L, \Gamma)$ 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 $l \subseteq L$.

To keep our definition simple, we assume that edges do not have any labels and that all relationships are mutual (undirected graph). From the social network represented as a graph it is possible to derive other data structures, representing aggregated information. The purpose of these data structures is to expose less information than the original form while enabling the computation of standard network metrics. Here, we define some of these structures:

- *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. It represents the neighborhood of v at distance 1.

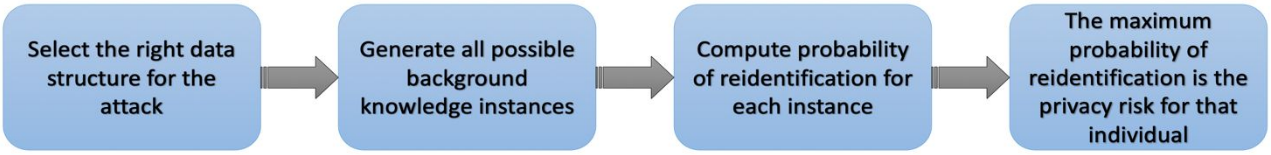
[1] The EU General Data Protection Regulation can be found at http://bit.ly/1TlgbjI.

- *Label vector* of an individual v is a set of labels $LA_v = \langle la_1 , la_2 ..., la_m \rangle$. Each \$$la_j$ = $(f, l)\, with\, j \in \{1, 2, ... |L|\}$ } is a pair of feature name f and label value l. The label vector of an individual can be empty. Each label describes a profile feature of an individual, such as gender or education etc.
- *Degree vector* of an individual v, denoted by $D_v = \langle d_{v_1}, d_{v_2}, \dots, d_{v_n} \rangle$, represents the number of neighbors of each neighbor of v. Thus, $d_{v_i} = len(F_{v_i})$.
- *Mutual Friendship vector* of an individual v, $MF_v = \langle mf_1, \dots, mf_n \rangle$ represents the number of common neighbors of v with each one of its neighbors v_i.

Given the definitions above, we can define a *Social Network Dataset* is a set of data structures $S = \{S_1, S_2, \dots, S_k\}$ where S_v $(1 \le v \le k)$ is the social network data structure of an individual v.

3. 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 by Chih-Hua Tai et al. [5] or Bin Zhou and Jian Pei [6]. In this paper we want to empirically study the privacy risk in social network data using the PRUDEnce risk assessment framework [4] PRUDEnce enables a privacy-aware ecosystem for sharing personal data. In this framework, Data Providers must perform privacy risk assessment before releasing the data to ensure the privacy of the individuals. The privacy risk assessment component of the framework produces a quantitative measure of privacy risk. Such measure depends on the kind of privacy attack simulated, the kind of data, and on the aggregation on the data itself. The simulation of a privacy attack requires two phases: first, we assume that a malicious adversary gathers, in some way, a *background knowledge* about an individual and then the adversary uses the acquired background knowledge to re-identify the individual in the social network dataset. A *background knowledge* is essentially a portion of an individual's data, used by the adversary to conduct an attack. It represents what kind of information the adversary has. The length of a background knowledge is the number of elements known by the adversary. Since, in principle, an adversary may have any kind of *background knowledge,* a worst-case scenario analysis has to be conducted. To this aim, PRUDEnce allows for the systematic generation of *background knowledge* through combinatorial generation: all possible *background knowledge instances* of a certain length are generated to simulate an attack, and the one *background knowledge instance* that present the highest probability of reidentification for an individual is the final privacy risk for that individual. The probability of re-identification is computed according to the principles of *k-anonymity* [7]. Thus, the more common the background of an individual is, the harder it is to re-identify him. The probability of re-identification depends then on the actual number of individuals that share the same *background knowledge instance* that the adversary possesses. In order to simulate how the adversary matches the background knowledge against the data, the attack must be formalized as a *matching function*, i.e. a function that indicates whether a *background knowledge instance* is present in the data of a certain individual. We can then compute the probability of re-identification as: $PR_S(s = v|b) = \frac{1}{|M(S,b)|}$ where $|M(S, b)|$ is $M(S, b) = \{s \in S \mid match(s, b) = True\}$, $s \in S$ is a record in the dataset, and b is a *background knowledge instance,* and *match* is the matching function. With this definition, privacy risk can vary between 0 (no risk) and 1 (maximum risk). The attack simulation process for each individual in the data, is summarized in Figure 1.



Select the right data structure for the attack

Generate all possible background knowledge instances

Compute probability of reidentification for each instance

The maximum probability of reidentification is the privacy risk for that individual

Figure 1: Attack simulation process

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. We define a set of attacks based on the data structures of a *Social Network Dataset* as previously defined.

- *Neighborhood Attack*: the adversary knows a certain number of friends/neighbors of an individual. It was introduced by Chih-Hua et al. [5]. The matching function is $match(b, F_v) = \{true\ if\ b \subseteq F_v\ false\ otherwise$
- *Label Pair Attack*: the adversary knows a certain number of pairs of features with their values of an individual. It was introduced by Chenyang Liu et al. [8]. The matching function is $match(b, LA_v) = \{true\ if\ b \subseteq LA_v\ false\ otherwise$
- *Friendship Degree Attack*: the adversary knows the degree of some neighbors of the victim and the degree of the victim. Introduced by Chih-Hua et al. [5]. The matching function is $match(b, D_v) = \{true\ if\ b \subseteq D_v\ false\ otherwise$
- *Mutual Friend Attack*: the adversary knows the number of mutual friends of the victim and some of its neighbors. Introduced by Chong-Jing et al [9]. The matching function is $match(b, MF_v) = \{true\ if\ b \subseteq MF_v\ false\ otherwise$

4. EMPIRICAL PRIVACY RISK ASSESSMENT

For experimental evaluation of the defined attacks, we use the Facebook Dataset provided by Stanford University's "Stanford Large Network Dataset Collection" [10]. This dataset includes node features (profiles), circles and ego networks. Nodes have been anonymized by replacing the Facebook-internal ids 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. Almost half of the all individuals have 30 friends/neighbors or less. In Figure 2 we can see some summary of the data.

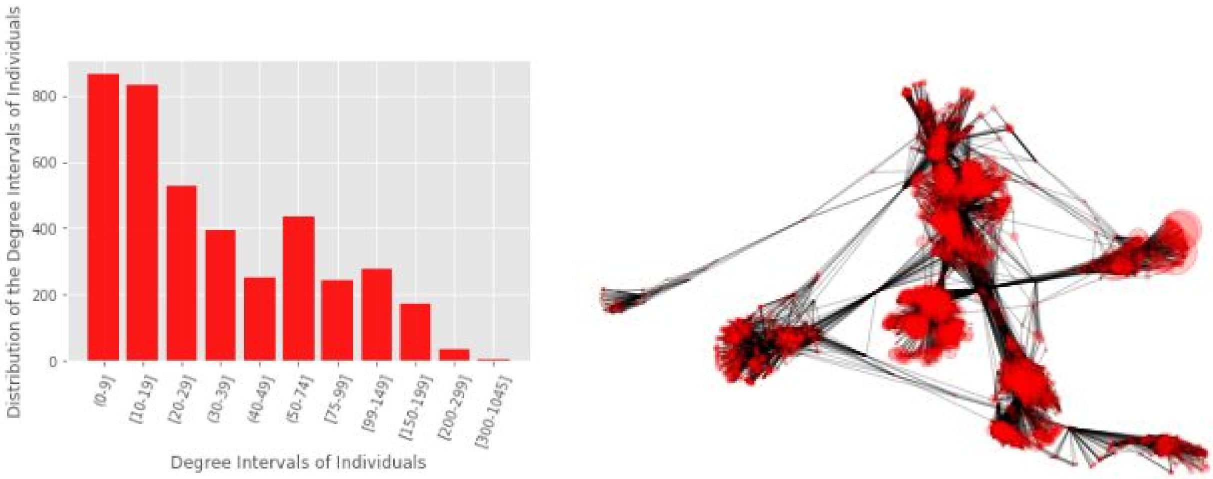


Figure 2 Degree distribution of nodes in the network and visualization of the Facebook Dataset

We simulated the privacy attacks defined in Section 3 with various background knowledge length (1,2,3 and 4) and computer the cumulative distribution of privacy risk, varying the length of background knowledge, for each attack. The cumulative distribution indicates on the y-axis the portion of individuals that present a privacy risk lower than the respective risk value on the x-axis. Therefore, the lower the curve, the higher the number of individuals with risk near 1. A higher curve instead indicates lower levels of risk overall. We present some of the results in Figure 3.

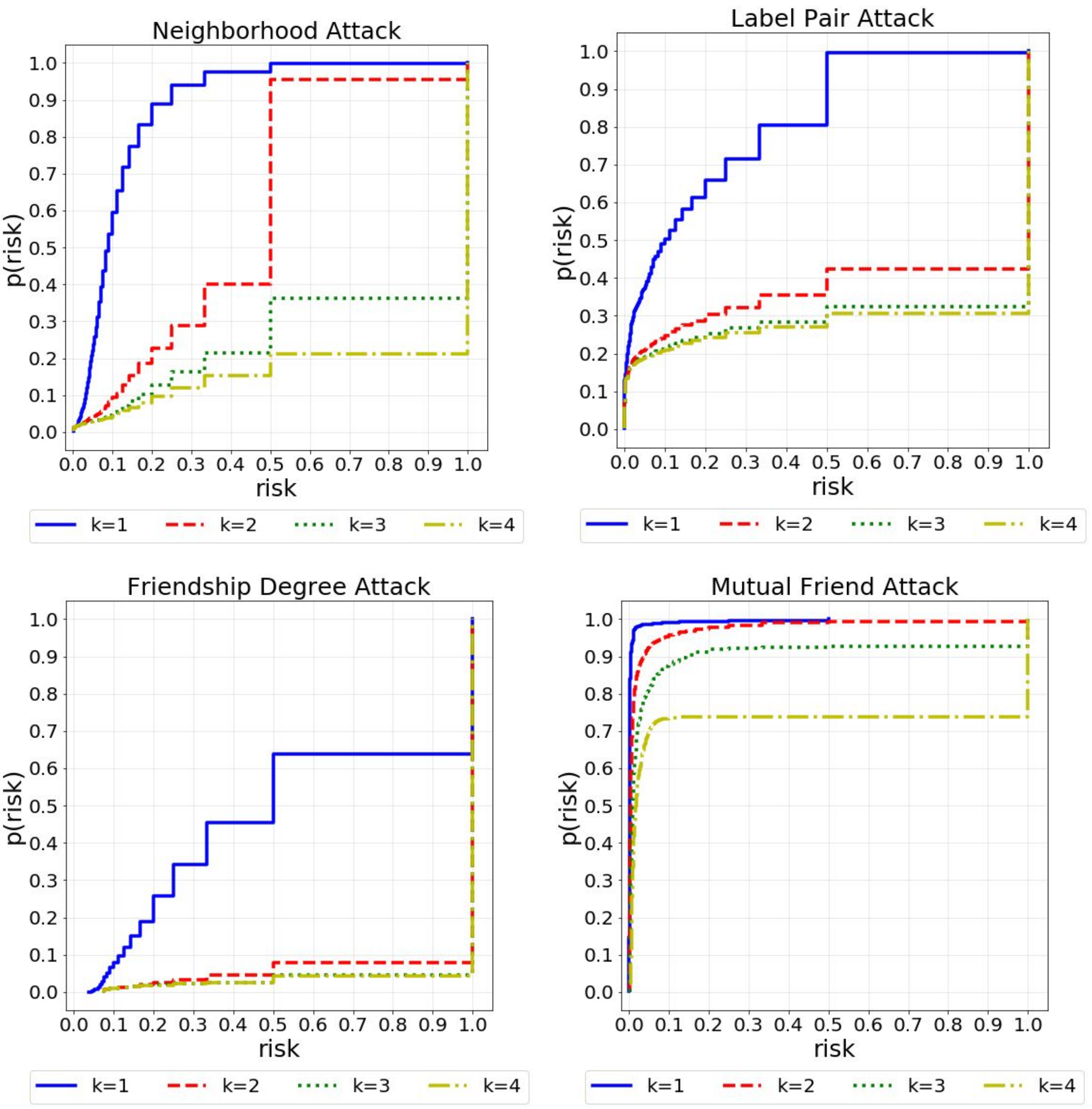


Figure 3 Cumulative distribution of risk for privacy attacks on network data

From Figure 3 we can see that, unsurprisingly, increasing the length of the background knowledge, i.e. increasing the amount of information that the adversary has regarding the victim, the privacy risk also increases. However, we see that, after the first increase from k=1 to k=2, privacy risk almost reaches a plateau. This behavior has been observed also in other contexts, for example in mobility data [11]. This suggests that, increasing the length of the background knowledge does not raise privacy risk proportionally. Also, interestingly, we see that the *Mutual Friend Attack* is the least powerful of all attacks, while the *Friendship Degree Attack* is the most dangerous. These results show that it is possible to analyze privacy risk empirically on network data, obtaining information that is easy to understand and to use in order to protect the data and the users in it.

BIBLIOGRAFY

[1] G. Rossetti, L. Milli, F. Giannotti and D. Pedreschi, "Forecasting success via early adoptions analysis: A data-driven study," *PloS one,* vol. 12, p. e0189096, 2017.

[2] G. Rossetti, L. Milli, S. Rinzivillo, A. Sıˆrbu, D. Pedreschi and F. Giannotti, "NDlib: a python library to model and analyze diffusion processes over complex networks," *Int. J. Data Sci. Anal.,* vol. 5, p. 61–79, 2018.

[3] L. Backstrom, C. Dwork and J. Kleinberg, "Wherefore Art Thou R3579x?: Anonymized Social Networks, Hidden Patterns, and Structural Steganography," in *Proceedings of the 16th International Conference on World Wide Web*, New York, NY, USA, 2007.

[4] F. Pratesi, A. Monreale, R. Trasarti, F. Giannotti, D. Pedreschi and T. Yanagihara, "PRUDEnce: a System for Assessing Privacy Risk vs Utility in Data Sharing Ecosystems," *Transactions on Data Privacy,* vol. 11, pp. 139-167, 2018.

[5] C.-H. Tai, P. S. Yu, D.-N. Yang and M.-S. Chen, "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, August 21-24, 2011*, 2011.

[6] B. Zhou and J. Pei, "Preserving Privacy in Social Networks Against Neighborhood Attacks," in *Proceedings of the 24th International Conference on Data Engineering, ICDE 2008, April 7-12, 2008, Cancún, Mexico*, 2008.

[7] L. Sweeney, "k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems,* vol. 10, p. 557–570, 10 2002.

[8] C. Liu, D. Yin, H. Li, W. Wang and W. Yang, "Preserving Privacy in Social Networks Against Label Pair Attacks," in *Wireless Algorithms, Systems, and Applications - 12th International Conference, WASA 2017, Guilin, China, June 19-21, 2017, Proceedings*, 2017.

[9] C.-J. Sun, P. S. Yu, X. Kong and Y. Fu, "Privacy Preserving Social Network Publication Against Mutual Friend Attacks," *Trans. Data Privacy,* vol. 7, p. 71–97, 2014.

[10] J. Leskovec and A. Krevl, *SNAP Datasets: Stanford Large Network Dataset Collection,* 2014.