Information Sciences 473 (2019) 87–100

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

On analyzing and evaluating privacy measures for social networks under active attack

Bhaskar DasGupta[a,*], Nasim Mobasheri[a], Ismael G. Yero[b]

[a] Department of Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA
[b] Departamento de Matemáticas, Escuela Politécnica Superior, Universidad de Cádiz, Algeciras 11202, Spain

ARTICLE INFO

Article history:
Received 2 May 2018
Revised 7 September 2018
Accepted 16 September 2018
Available online 18 September 2018

2010 MSC:
68Q25
68W25
05C85

Keywords:
Privacy measure
Social networks
Active attack
Empirical evaluation

ABSTRACT

Widespread usage of complex interconnected social networks such as *Facebook*, *Twitter* and *LinkedIn* in modern internet era has also unfortunately opened the door for privacy violation of users of such networks by malicious entities. In this article we investigate, both theoretically and empirically, privacy violation measures of large networks under active attacks that was recently introduced in Trujillo-Rasua and Yero (2016). Our theoretical result indicates that the network manager responsible for prevention of privacy violation must be very careful in designing the network *if its topology does not contain a cycle*. Our empirical results shed light on privacy violation properties of eight real social networks as well as a large number of synthetic networks generated by both the classical Erdös–Rényi model and the scale-free random networks generated by the Barábasi–Albert preferential-attachment model.

1. Introduction

Due to a significant growth of applications of graph-theoretic methods to the field of social sciences in recent days, it is by now a standard practice to use the concepts and terminologies of network science to those social networks that focus on interconnections between people. However, social networks in general may represent much more than just networks of interconnections between people. Rapid evolution of popular social networks such as *Facebook*, *Twitter* and *LinkedIn* have rendered modern society heavily dependent on such virtual platforms for their day-to-day operation. The powers and implications of social network analysis are indeed *indisputable*; for example, such analysis may uncover previously unknown knowledge on community-based involvements, media usages and individual engagements. However, all these benefits are *not* necessarily cost-free since a malicious individual could compromise privacy of users of these social networks for harmful purposes that may result in the disclosure of sensitive data (attributes) that may be linked to its users, such as node degrees, inter-node distances or network connectivity. A natural way to avoid this consists of an "anonymization process" of the relevant social network in question. However, since such anonymization processes may *not* always succeed, an important research goal is to be able to quantify and measure how much privacy a given social network can achieve. Towards this goal, the recent work in [42] aimed at evaluating the *resistance* of a social network against active privacy-violating attacks

* Corresponding author.
E-mail addresses: bdasgup@uic.edu (B. DasGupta), nmobas2@uic.edu (N. Mobasheri), ismael.gonzalez@uca.es (I.G. Yero).

https://doi.org/10.1016/j.ins.2018.09.023
0020-0255/

Table 1
List of real social networks studied in this paper.

Name	# of		Description
	nodes	edges	
(A) Zachary Karate Club [47]	34	78	Network of friendships between 34 members of a karate club at a US university in the 1970s
(B) San Juan Community [31]	75	144	Network for visiting relations between families living in farms in the neighborhood San Juan Sur, Costa Rica, 1948
(C) Jazz Musician Network [22]	198	2842	A social network of Jazz musicians
(D) University Rovira i Virgili emails [23]	1133	10903	The network of e-mail interchanges between members of the University Rovira i Virgili
(E) Enron Email Data set [15]	1088	1767	Enron email network
(F) Email Eu core [36]	986	24989	Emails from a large European research institution
(G) UC Irvine College Message platform [37]	1896	59835	Messages on a Facebook-like platform at UC-Irvine
(H) Hamsterster friendships [24]	1788	12476	This Network contains friendships between users of the website hamsterster.com

by introducing and studying theoretically a new and meaningful privacy measure for social networks. This privacy measure arises from the concept of the so-called k-metric antidimension of graphs that we explain next.

Given a connected simple graph $G = (V, E)$, and an ordered sequence of nodes $S = (v_1, \ldots, v_t)$, the *metric representation* of a node u that is *not* in S with respect to S is the vector (of t components) $\mathbf{d}_{u,-S} = (\text{dist}_{u,v_1}, \ldots, \text{dist}_{u,v_t})$, where $\text{dist}_{u,v}$ represents the length of a shortest path between nodes u and v. The set S is then a *k-antiresolving set* if k is the largest positive integer such that for every node v not in S there also exist *at least* other $k-1$ different nodes $v_{j_1}, \ldots, v_{j_{k-1}}$ not in S such that $v, v_{j_1}, \ldots, v_{j_{k-1}}$ have the *same* metric representation with respect to S (*i.e.*, $\mathbf{d}_{v,-S} = \mathbf{d}_{v_{j_1},-S} = \cdots = \mathbf{d}_{v_{j_{k-1}},-S}$). The *$k$-metric antidimension* of G is defined to be value of the minimum cardinality among all the k-antiresolving sets of G [42]. If a set of attacker nodes S represents a k-antiresolving set in a graph G, then an adversary controlling the nodes in S cannot *uniquely* re-identify other nodes in the network (*based on the metric representation*) with probability higher than $1/k$. However, given that S is unknown, any privacy measure for a social network should quantify over *all* possible subsets S of nodes. *In this sense, a social network G meets (k, ℓ)-anonymity with respect to active attacks to its privacy if k is the smallest positive integer such that the k-metric antidimension of G is no more than ℓ. In this definition of (k, ℓ)-anonymity the parameter k is used for a privacy threshold, while the parameter ℓ represents an upper bound on the expected number of attacker nodes in the network.* Since attacker nodes are in general difficult to inject without being detected, the value ℓ could be estimated based on some statistical analysis of other known networks. A simple example that explains the role of k and ℓ to readers is as follows. Consider a complete network K_n on n nodes in which every node is connected with every other node. It is readily seen that for any $0 < \ell < n$, this network meets $(n-\ell, \ell)$ -anonymity. In other words, this means that a social network K_n guarantees that a user cannot be re-identified (based on the metric representation) with a probability higher than $1/(n-\ell)$ by an adversary controlling at most ℓ attacker nodes. For other related concepts for metric dimension of graphs, the reader may consult references such as [14,25,29].

Chatterjee et al. [9] (see also [48]) formalized and analyzed the computational complexities of several optimization problems motivated by the (k, ℓ)-anonymity of a network as described in [42]. In this article, we consider three of these optimization problems from [9], namely Problems 1–3 as defined in Section 2. A high-level itemized overview of the contribution of this article is as follows (see Section 3 for precise technical statements and details of all contributions):

- ▷ Our theoretical result concerning the anonymity issues for networks without cycles is provided in Theorem 1 in Section 3.1. Some consequences of this theorem are also discussed *immediately following a statement of the theorem*.
- ▷ In Section 3.2, we first describe briefly efficient implementations of the high-level algorithms of Chatterjee et al. [9] for Problems 1–3 (namely Algorithms 1 and 2 in Section 3.2.1). We then tabulate and discuss the results of applying these implemented algorithms for the following type of network data:
  - ▷ eight real social networks listed in Table 1 in Section 3.4.2,
  - ▷ the classical undirected Erdös–Rényi random networks $G(n, p)$ for four suitable combinations of n and p, and
  - ▷ the *scale-free random networks* $G(n, q)$ generated by the Barábasi–Albert *preferential-attachment* model for four suitable combinations of n and q.

  The 6 tables that provide tabulations of the empirical results are Tables 2–7 in Section 3.2 and the type of conclusions that one can draw from these tables are stated in the 11 conclusions numbered ①–⑪ in the same section. Despite our best efforts, we do not know of any other alternate approaches (*e.g.*, sybil attack framework) that will provide a significantly simpler theoretical framework to reach all the 11 conclusions as mentioned above.

As an illustration of a potential application, consider the *hub fingerprint query* model of Hey et al. [26]. Noting that the largest hub fingerprint for a target node u is the metric representation of u with respect to the hub nodes, results on (k, ℓ)-anonymity are directly applicable to this setting of Hey et al. [26] that models an adversary trying to identify the hub