




Think Before You Share in OSNs: Textual Content and Connection Weight Put You at Higher Privacy Risk

Önder Çoban¹, Ali İnan², Selma Ayşe Özel³

¹Adıyaman University, Computer Engineering Department, Adıyaman, Turkey

²Adana Alparslan Türkeş Science and Technology University, Computer Engineering Department, Adana, Turkey

³Çukurova University, Computer Engineering Department, Adana, Turkey

Corresponding Author: onder.cbn@gmail.com

Research Paper

Received: 10.03.2022

Revised: 11.04.2022

Accepted: 19.04.2022

Abstract—The widespread use of OSNs has brought forward the issue of privacy protection over OSNs, as sensitive information of users needs to remain private. Most users are unaware of possible privacy risks associated with sharing personal information in their accounts. Privacy settings of OSNs focus on protecting users' information just by providing them with means of configuring the audience of shared information. As such, privacy risk estimation (or scoring) is a hot topic in the field of OSN research and aims to develop risk measuring tools to ensure user privacy in OSNs. Conventional studies in the area often rely on synthetically generated or survey-based data and do not make any effort to infer private attribute values of users to utilize inference success in privacy scoring of these users. In this study, we propose a novel framework that involves populating a response matrix by using attribute inference and obtaining network aware-risk scores not just by using users' connections but weights of these connections as well. We perform attribute inference of users based on both their textual contents and connections. Our rule-based inference mechanism employed on contents produces inference accuracies ranging from 0.54 to 1.0 depending on the attribute at hand. On the other hand, the inference mechanism involving users' social connections produces inference accuracies of 1.0 almost for all of the considered attributes. We present results and challenges of attribute inference and use inferred attributes in privacy risk scoring. In addition, unlike existing works, we use and show that social tie strengths have to be taken into account in network-aware privacy risk scoring.

Keywords—Facebook, privacy risk estimation, social tie strength, online social networks.

1. Introduction

A huge number of people in today's world have moved their social interactions to Online Social

Networks (OSNs) such as Facebook, Twitter, and LinkedIn [1]–[3]. These OSNs paved way for users to have a digital representation (i.e., profile) and various opportunities such as sharing data of differ-

ent kinds, learning new things about their hobbies, getting news, building and maintaining connections as well as their offline connections [4], [5]. OSN profiles include both structured and unstructured data of users [6]. Popular OSNs directly or indirectly demand users to share increasingly more data - of even a sensitive nature - so that other users can find them more easily [1], [3], [7]–[9]. Consequently, OSN users often willingly disclose their personal information and they are not usually aware of who will or will not have access to what they have just published [3], [9]. While dissemination of information in the real world is slow and almost local, information shared publicly over OSNs can be retrieved on the Internet anytime, anywhere, and by anyone [10]. Additionally, the private information of OSN users can also be effectively inferred by using different techniques including graph analysis and probabilistic classification techniques [4], [11]. For instance, according to [12], it is possible to estimate some private traits of a user's personality by leveraging Facebook activities. Another study states that it is even possible to infer user characteristics from the attributes of users who are part of the same community [13]. On the other hand, third-party OSN applications and befriended users (i.e., users cannot analyze their friends' time-varying and changing behavior) may also put the information owner at risk in OSNs [14].

As a result, any information shared on OSNs may be sensitive and any kind of disclosure can significantly harm users' privacy [2] by giving way to some threats such as identity theft, digital stalking, building consumer models for advertising, identity cloning, and blackmailing among many others [1], [14]–[16]. As such, both users and OSN service providers recognize the need to ensure that user privacy is well preserved in OSNs [3], [5]. OSN providers have considerably improved their privacy

protection tools by introducing more controls like groups, lists, and circles [7], [10], [17]. However, at the end of the day, the most powerful data protectors are the users themselves [18], as a user's level of privacy attitude determines her sharing activities on any OSN. The privacy-aware user tends not to share her own or her friends' sensitive or private information, whereas an unaware user does not carry such a concern. Access control policies enforced by OSN providers are based on users' privacy settings that are confusing and time-consuming to manage. These policies make users face too many options and controls, which cause a lack of understanding of the privacy risks and threats or being unable to accurately assess them [3], [5], [9], [19]–[21]. Thus, determining privacy risks when publishing information on OSNs often presents a challenge for the users.

A measure of how much sensitive information users shared with others on an OSN would help the users understand whether they individually share too much [3]. However, privacy measurement and quantification is a challenging task because privacy does not have a specific definition as its degree differs from individual to individual. In other words, there are different perceptions and concerns of privacy among different cultures, notions, societies, and religions [16]. Even so, there has been a large collection of studies with a different point of view that considers privacy in OSNs.

Conventional privacy-related studies mainly focus on protecting or measuring the identities or private attributes of users [22]. On the other hand, content-based measuring has not been strongly addressed before and existing privacy models for structured data are inadequate to capture privacy risks from user posts [23]. However, a lot of users' personal aspects can be extracted from user-generated contents [24], as a text message or post may directly or

indirectly disclose personal information like a user's gender, age, political view, interests, and hobbies that many users would not prefer to reveal directly [23], [25]. In the literature, few studies [22]–[24] consider content in privacy scoring, while most of them do not include methods for obtaining/driving/inferring sensitive information from the messages and status updates [2], [17], [26].

Content-based sensitivity detection is studied in some works, but these studies do not address the problem in terms of privacy measurement [27], [28]. Instead, these works often focus on the detection of sensitive contents. Additionally, they also try to categorize contents into more general private information (i.e., vacation, location, health, etc.) or personal attributes (i.e., gender, political view, etc.) [27], [28]. All of these cases show that the majority of existing privacy scoring solutions are inspired from [1] and score users based on the attribute they reveal to other users.

In this paper, we propose and implement a novel privacy risk scoring framework that relies on attribute inference from both content and social connections of users. This framework populates a response matrix using disclosed and inferred attributes in an incremental way and produces network-aware risk scores of users. Computing network-aware risk scores involve calculation of intrinsic risk scores at first and then giving them along with social tie strengths to produce the final output.

Our approach is described herein for one of the most popular OSNs, namely Facebook, but for obvious reasons, applies to any other OSN trivially. Our contributions in this paper can be summarized as follows:

- Unlike most of the existing studies, we perform privacy scoring over real-world Facebook data.
- We use different privacy risk scoring methods and explore correlations between these meth-

ods.

- We build various adversarial models with different attribute inference capabilities based on content and structure information of OSN users and show that users are at considerably higher risk against these adversaries compared to their risk against a passive adversary.
- We adopt the self-information feature weighting method to measure the strength of the social tie across OSN users in a non-symmetrical way.
- We show that the strength of the social tie from one user to another plays a statistically significant role in computing centrality-based (or network-aware) risk scores of OSN users.

This paper is an extension of our preliminary work [29] and to the best of our knowledge, it is the first study to concentrate on the privacy scoring of Facebook users from Turkey. This paper also suggests using social tie strength in network-aware risk calculation. It also performs and reports attribute inference based on users' textual contents and network structure. We believe that these make this study different from the existing studies in the field of privacy scoring over OSNs.

2. Literature Review: Privacy Evaluation And Scoring In OSNs

Either consciously or unconsciously, OSN users disclose on their profiles sensitive information that puts their privacy at risk. Several techniques and methods that aim at quantifying the extent of such risks have been proposed. The table given in the appendix presents a summary of the studies with a focus on privacy risk scoring over OSNs. As seen from the table, previous studies can be grouped into two categories based on whether they perform PRE (privacy risk estimation) or suggest a novel framework or approach for privacy evaluation of

OSN users.

In the first category, one of the earliest and most popular works based on profile items belongs to Liu and Terzi [1] who have proposed an Item Response Theory (IRT) based model to assess the privacy risk of OSN users. It also has been a guide for later studies [2], [3], [10], [20], [26], [30]–[37] which propose a new or use an existing PRE method based on this work. Privacy-functionality score (PFS) [34] is one of these later studies that measure privacy by dividing the amount of information a user can see about other users by the amount of information it shows about herself. There are other methods or metrics [9], [15], [18], [19], [22], [38], [39] as well, which have a focus on privacy evaluation in OSNs. For instance, Susceptible-Infectious-Recovered (SIR) model is adopted for modeling the spread of information in an OSN in [18].

In the second category, on the other hand, studies [5], [7], [11], [14], [17], [21], [23]–[25], [40]–[42] focus on building a new and complete framework or model for PRE and increasing privacy awareness in OSNs. The majority of these studies suggest using three types of information, but they often do not concern with the techniques to be employed. For instance, SONET defines attribute to actor, attribute to attribute, and actor to actor relationships and contains a privacy index (PIDX) based on the visibility and sensitivity of user attributes. Besides, there are developed OSN applications [4], [43] as well, which mostly aim to increase the privacy awareness of users. For instance, Friend Inspector is a serious game that allows its users to playfully increase their privacy awareness on Facebook.

When we compare the previous studies in terms of their main focus, it is clear that the majority of them consider risk scoring of users, while some of them take privacy in different aspects such as

measuring the extent of information diffusion [9], [18], privacy setting configuration [21], predicting privacy vulnerability of users [44], sharing behavior analysis [45], [46], modeling of privacy attacks [8], and privacy-preserving friending [6], [47]. Similarly, the majority of the works consider users' point of view when measuring privacy risk, while a few are service oriented [14] and consider third-party applications [11], [36].

As a branch of link prediction, tie strength prediction has also received much attention from researchers in recent years. In the literature, tie strength prediction is one of the hot topics in the OSN research field and many studies try to achieve this task using different techniques such as supervised learning [48], [49], regressions models [50], similarity-based models [51], simple partial [52] or linear functions [49], [53] and so on.

On the other hand, when the previous studies are examined in terms of the type of information that they use, we observed that most studies rely on structured data (i.e., profile items) and a few suggest employing actions and content in the privacy risk scoring method or model. We also observe that content-based studies are typically not concerned with the technologies that can extract or infer attribute values from the unstructured data [2], [38]. It is important to note that some content-based studies aim at detecting sensitive information from OSN data without any emphasis on privacy evaluation and scoring. As of the writing of this paper, we were able to detect only a handful of studies [22]–[25] that use content in the privacy evaluation of OSN users.

In this study, however, we studied the attribute inference of users based on their profile items, textual contents, and connections. Afterward, inferred attributes along with the disclosed ones are used for the privacy scoring of users. Privacy risk scoring is

Table 1.
 Quantitative description of crawled public Facebook data [45].

Property (# of)	Count	Property (# of)	Count
users (nodes) whose entire profile has been crawled	20,000	public wall	18,579
friendship links (edges) among crawled users	402,300	private wall	1,421
friendship links	3,980,270	private friend list	11,675
public and partially public friend list	8,325	liked pages [54]	459,335
unique accounts discovered	2,350,454	wall activities (post, comment, reply)	5,972,531
users disclosing date of birth	1,687	users disclosing age gender	16,286
users whose age is detected/computed	879	users disclosing places	11,113

performed by using both profile attributes and social connections (i.e., network structure) of users. In network-structure-based scoring, social tie strengths of users are taken into consideration as well.

We believe that content-based attribute inference and using social tie strengths in network-aware risk scoring make this study different from the existing ones.

3. Material

3.1. Data set

The data sets utilized in this study were collected with a web crawler that obtains public Facebook data from user accounts. The design choices and implementation details of the crawler, along with detailed statistical analyses of the collected OSN data can be found in both [45] and [55]. This crawler outputs a breadth-first traversal/search (BFS) of all public, Turkish Facebook accounts starting at a seed account. It then visits the seed user’s friends, friends of the seed’s friends, and so on. Thus, crawling creates a process similar to the waves of a water droplet.

In this phase, the crawler visits an account and fetches all public profile information, wall content,

and friend list of the current account. The crawler then stores the collected friend list into a database so as to visit the accounts within the list in the next steps of crawling and this process continues until it is stopped. The only criterion in the crawling process is that the crawler just collected every public information from user accounts in text format. That is, multimedia and other non-textual information is discarded. Additionally, information about groups that users joined and pages that users liked are not crawled.

We would like to note that the crawling process is started by using just a few accounts that were befriended with the seed account. No additional criterion is applied to decide how to select accounts to visit by the crawler. The crawling process is performed in BFS order and therefore accounts befriended with the first, second, and any other levels of friends are found to be potential (since the crawling process was stopped after visiting 20K users, so remaining unvisited ones can be referred to as potential) candidates to be visited by the crawler.

In this study, we work with the largest sub-graph generated by the crawler of [45]. The statistical properties of this graph are summarized in Table 1. The graph contains basic profile information of 20K users from Turkey, alongside their friendship links

Table 2.
 List of attributes used in the privacy risk estimation process.

Attribute/Item	Code	Attribute/Item	Code	Attribute/Item	Code	Attribute/Item	Code
Friend List	FL	Political View	PV	Email	MA	Having Child Status	CHL
Gender	G	School	EDU	Marriage Date	MD	Relationship Status	REL
Birth Date	BD	Place (Hometown)	HT	OSN accounts	CON	Family Membership	KIN
Age	AG	Place (Lived-in)	LIV	Places (Other)	OP	Phone Number	PN

Table 3.
 A sample list of Facebook pages associated with political parties.

Party		Facebook Page		Party		Facebook Page	
Name	Code	ID	Title	Name	Code	ID	Title
Justice and Development Party	JDP	4*****1	T C Devlet Başkanı Erdoğan	Good Party	GP	1*****4	Meral Akşener
		5*****7	Efsane Lider R T Erdoğan			1*****6	Cumhurbaşkanı Adayım Meral Akşener
		7*****6	AkParti			2*****5	İYİ Parti Gönüllüleri
Nationalist Movement Party	NMP	7*****3	Lider Devlet Bahçeli	Peoples' Democratic Party	DPP	5*****3	Selahattin Demirtaş
		7*****3	Ülkücü Hareket Engellenemez			1*****4	Sırrı Süreyya ÖNDER HDP
		4*****0	Milliyetçi Hareket Partisi MHP			1*****2	HDP Lideri Selahattin Demirtaş
Republican People's Party	RPP	4*****9	Kemal Kılıçdaroğlu	Felicity Party	HP	2*****8	Temel Karamollaoğlu
		1*****7	Cumhuriyet Halk Partisi			1*****4	Saadet Partisi Milli Görüş
		1*****4	Ne Mutlu CHP liyiz			9*****4	Saadet Partisi Pendik Gençlik Kolları Resmi Sayfası

(approx. 4M), and wall activities (over 5.9M) which are comprised of posts, comments, and replies. Notice that liked pages of users are collected under a different study aiming at proposing a new privacy scoring framework [54].

As seen in Table 1, the largest snapshot of the crawled data contains a high volume of user data and is sufficient for making rule-based attribute inference. Please note that even in the case of automatic inference, the data is very sufficient to create large datasets to train learning models as done in [56]. In this study, we disregard the underlying graph structure and focus only on basic profile information and 5.9M wall activities to perform

privacy risk scoring of Facebook users. Notice that this largest snapshot is a graph with 20K nodes and 402,3K directed edges. To perform attribute inference and PRE, we consider 16 different personal attributes of users. These attributes are given along with their codes in Table 2. The reader is advised to [55] for all of the details of the crawling process as well as the detailed statistical description of the crawled data with respect to its different snapshots.

3.2. Lexicon of political pages

This lexicon includes a list of Facebook pages along with their supported political party. To create this lexicon, we started with a subset of 459,335

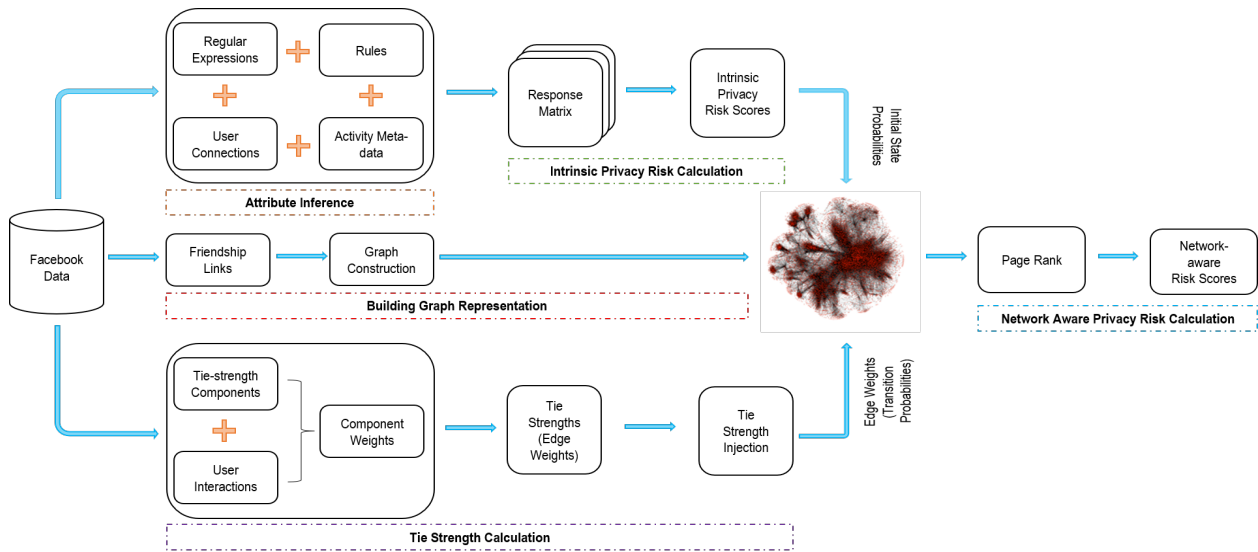


Figure 1. Flowchart of our privacy risk scoring method.

Facebook pages liked by 13,023 of 20K users in the largest snapshot of the crawled data. As stated before, our crawler does not traverse liked pages of users, but this additional data is crawled by Kılıç and Inan [54] in their separate study that proposes a quantitative framework for privacy risk score evaluation. Hereby, we could create a subset of liked pages that are created for political purposes.

After creating our subset, we selected the pages whose title includes any term related to six main political parties (i.e., JDP, NMP, RPP, HP, DPP, and GP) in Turkey. For instance, if the title of a page includes “Devlet Bahçeli”, we assigned the related page to the pages supporting the NMP. Similarly, if the title of a page contains “Kemal Kılıçdaroğlu”, we assigned the related page to pages supporting the RRP party. This process produced a list including 1,763 pages such that each page is associated with one of the six parties.

Table 3 gives an example and incomplete list of pages with their associated political parties. We used

this lexicon to infer the political views of users based on their liked pages (see Section 4.1.2)

3.3. Lexicon of schools and universities

This lexicon contains the names of primary, secondary, and high schools [57], [58] in Turkey. It also includes university names in Turkey along with the names of cities they are located [59].

3.4. Lexicon of district and provinces

This lexicon includes province and district names in Turkey. Each record in this lexicon stores information on the province a district is located at.

4. Methods

This section describes the methods we used to perform privacy risk scoring of Facebook users. Figure 1 depicts a flowchart of our privacy risk scoring framework that includes the following steps:

- Attribute inference: In this step, we perform attribute inference and build different response

matrices to explore the effect of attribute inference on both intrinsic and network-aware privacy risk scores of users.

- Building graph representation: Converts relational Facebook data into a graph using friendship links among users.
- Tie strength calculation: Calculates tie strength between each pair of befriended users based on our tie strength components. Computed tie strengths are incorporated into the graph as edge weights (i.e., transition probabilities).
- Intrinsic privacy risk calculation: Computes intrinsic privacy risk scores based on sensitivity and visibility components obtained from the response matrix. These intrinsic risk values are incorporated into the graph as relative importance (i.e., initial state value) of nodes during the traversal of nodes (i.e., users).
- Network-aware privacy risk calculation: Performs privacy risk measurement by running a personalized page rank algorithm on the final graph resulting from the previous steps briefly explained above.

The natural choice of a data structure to represent OSNs is a graph. As such, we consider a set of n users participating in an OSN as a directed graph $G(V, E)$, where V is the set of nodes $\{v_1, v_2, \dots, v_n\}$ in which each node $v_j \in V$ represents a user, and $E \subseteq V \times V$ is a set of directed edges $\{e_1, e_2, \dots, e_m\}$ such that an edge $e_q \in E$ represents a friendship link between a pair of users $v_j, v_{j'} \in V$. Each user in V may disclose information related to a set of t topics $T = \{t_1, t_2, \dots, t_z\}$ that corresponds to his/her profile items such as gender, age, job, religion, and so on [31]. Following subsections present the details of our proposed method.

4.1. Building a Response Matrix

A response matrix R is an $n \times z$ matrix which is associated with the set V of users and the set T of attributes [31]. Response matrix R includes attributes disclosed by users in the *dichotomous* and *polytomous* cases. In the *dichotomous* case, cells of R take values in $\{0, 1\}$, while in the *polytomous* case individual cell values are integers in range $\{0, 1, \dots, l\}$ [1]. In a *dichotomous* R , entry $r_{ji} = 0$ implies that user v_j keeps item t_i private, whereas $r_{ji} = 1$ represents that user v_j makes item t_i publicly available. In a *polytomous* R , $r_{ji} = 0$ has the same meaning compared with the *dichotomous* case, while $r_{ji} = k$ with $k \geq 1$ means that user v_j discloses information regarding item t_i to users which are at most k links away in G [1], [31].

We make the critical observation that in addition to which attributes a user reveals willingly, the attributes that can be inferred successfully by an adversary should be incorporated into the process of risk scoring. Hence, a deeper and more effective privacy measurement of OSN users can be achieved. As a side benefit, users could be informed that they may be under considerable privacy violation risk even if they keep their data private. To do so, in this paper, we use three different models to build our response matrix. These models are described in the following sub-headings.

4.1.1 L1: Revealed Attributes

This model considers only revealed attribute values (e.g., similar setting as in [1] and most of the related work) to build a response matrix.

4.1.2 L2: Attributes Inferred from Contents

Uses meta-data and contents of all activities of a user's wall together with the profile items and

liked pages to infer private attribute values. This model utilizes regular expressions, some rules, and activity meta-data in the inference phase. We group our attributes in Table 2 into 6 groups based on the inference mechanism as follows:

- G1: This group includes attributes MA, PN, and CON. The inference mechanism utilizes regular expressions. For instance, we search URLs in attribute fields and the activity contents of a user. Next, we filter URLs to select ones that include OSN names (e.g., Twitter, Instagram, etc.) and extract the username from these connection URLs. Similarly, we search email addresses and phone numbers within the activity contents and attribute fields of users.
- G2: This group includes attributes REL and CHL. To infer these two attributes, we first seek for relationship and family membership fields of both the account owner and his/her friends' profiles to check whether they have disclosed a private relationship or family membership with the wall owner. If no information is found, a rule-based inference mechanism is employed to infer the target attribute values from the wall activities.

Let V_A be a Facebook user whose REL and CHL attributes will be inferred, while $F = \{v_1, v_2, \dots, v_f\}$ represents a set of befriended users with V_A . Additionally, let W_A represents a set of activity contents written by the account owner (i.e., V_A), while W_F stands for a set of activity contents both addressing (i.e., targeting) user V_A and written by any of his/her friends in F . To infer related attributes of user V_A , we search all textual content in both W_A and W_F for a phrase p in a set of phrases denoted by P that includes single or multi-word phrases specific to the target attribute. If any $p \in P$ is observed in $W_A \cup W_F$, it is assumed that the target attribute

is inferred by taking its orientation into account. For the REL attribute, P contains phrases that strongly imply that V_A 's relationship status has changed. These phrases are selected based on our observations and sample phrases are as follows: “*düğünüm var (I have a wedding ...)*”, “*artık evliyim (I am married now)*”, “*evlendik (we got married)*”, “*düğünüme bekliyorum (waiting for you to attend my wedding)*”, “*evlilik yıl dönümü (wedding anniversary)*”, “*eşimle beraber (together with my spouse)*”, “*mutluluklar (I wish you to be happy)*”, “*ömür boyu mutlu (be happy for life)*”, “*bir yastıkta (be happy for life)*”, “*güzel gelinim (my beautiful bride)*”, “*eşinle birlikte (together with your spouse)*”, and “*mutlu mesud (I wish you to be happy)*”.

For the CHL attribute, the same process is employed on P including phrases that imply that user V_A has a child. Sample phrases are as follows: “*analı babalı.. (wish your child grow up with parents..)*”, “*bebeğin çok tatlı (your baby is very sweet)*”, “*yanaklarını yerim (I eat your cheeks)*”, “*allah başı şulasın (god bless)*”, and so on.

- G3: This group contains attributes FL, BD, AG, MD, and EDU. Inference of these attribute values relies mainly on wall activities, user interaction listed under wall activities, and meta-data of activities.

To infer the friend list of a wall owner, we search users who interact (i.e., posting a status, commenting on the wall owner's post, etc.) with the wall owner on both their own wall pages and wall owner's wall.

On the other hand, the simple rule-based inference mechanism employed to infer attributes in G2 is again employed to infer the BD, AG, EDU, and MD attribute values in a similar way. In this phase, we only consider event and direct

post within $W_A \cup W_F$. To infer BD, we create P to include some multi-word phrases which give clues about the birth day of V_A , while we add phrases into the P implying educational changes, marriage, and birthday celebration to infer EDU, MD, and AG attributes respectively. For the BD attribute some of phrases in P are as follows: “*doğum günü (birthday..)*”, “*nice yaşlara (happy birthday)*”, “*tarihinde doğdu (born on..)*”.

For the EDU attribute, sample phrases are as follows: “*okula başladı (started school)*”, “*okulu terk etti (dropped out of school)*”, “*mezun oldu (graduated)*”.

For the MD attribute, sample phrases used in P are as follows: “*düğünüm var (I have wedding)*”, “*evlendim (I got married)*”, and so on. Note that if any of $p \in P$ observed in just event and direct posts within $W_A \cup W_F$, we use the orientation of the p to infer the EDU attribute, while we use the most frequently observed posting date to infer the BD attribute of V_A . To infer the AG attribute, on the other hand, we only consider event posts that show the exact date of birth information. Finally, we infer the MD attribute by taking the posting date of activities.

If any p is not observed in event posts, we additionally use our lexicon of schools and universities (see Section 3.3) and try to match any of the school names within any part of the activity contents within $W_A \cup W_F$.

- G4: Inference is quite similar to the mechanism employed to infer attributes in both G2 and G3. This group is comprised of attributes HT, LIV, and OP. To infer these attributes, the inference mechanism applies lexicon-based searching over all textual content (i.e., W_A) generated by the wall/account owner (i.e., V_A). To infer these three attributes, we use a com-

mon lexicon (see Section 3.4) that includes district and province names in Turkey. However, we apply simple rules to differentiate (i.e., hometown or live-in place) the values of these attributes.

If we find a place name in W_A , we also search for phrases P in both the preceding and following words of the place name. If any p in P exists before or after the place name, we add this place as one of the possible candidates of HT or LIV attribute. For the LIV attribute, we additionally consider posting location information, if exists, to infer possible candidates of users' places. For the OP attribute, only place names matched with the lexicon are considered to be candidate places.

For the HT attribute P includes the some of the followings: “*baba ocağı (father's home)*”, “*memleketim (my hometown)*”, “*doğduğum yer (where I was born)*”, and “*bizim köy (our village)*”.

For the LIV attribute, sample phrases in P are as follows: “*ilk iş günü (first business day)*”, “*iş yerim (my working place)*”, “*okulum (my school)*”, “*taşındım (moved my house)*”, and “*yeni evim (my new house)*”.

- G5: This group includes only the attribute PV. The inference of this attribute is based on the users' liked pages. To achieve this task, we used our dictionary of political pages (see Section 3.2) and inferred which party is supported by the wall owner. At this stage, we count the number of liked political pages by the wall owner, and the PV value is determined by looking at which of the six main parties (i.e., JDP, NMP, RPP, HP, DPP, and GP) was liked more frequently.
- G6: This group includes only the attribute G. We employ an inference mechanism that builds a learning model for determining the gender

of the wall owner. The details of this learning model are presented in [56].

4.1.3 L3: Attributes Inferred from Social Connections

This model considers inference on social connections of the users and mainly depends on the bi-directional structure of Facebook. Contrary to the L2 model, it is not possible to employ attribute inference mechanisms to determine the values of all attributes that can be obtained by using the social connections of users. This is because the data used in this study is real data and it is incomplete. As such, in the L3 model, we only consider a few attributes that we could be able to obtain, and group them into 3 groups as follows:

- G1: This group includes the FL attribute. We extract the friend list of a user by performing reverse social engineering over the OSN [56], [60].
- G2: This group includes only the attribute G. Inference mechanism is based on the popularity of the user's first name among other users in the crawled snapshot [56].
- G3: This group contains the attributes KIN, REL, and CHL. The inference mechanism again depends on the bi-directional nature of Facebook. For relationship status (i.e., REL), we only search for any user who discloses that he/she is in a relationship with the related user. For family membership attributes, on the other hand, we infer all kinds of family memberships of the related user if each of these relationships is disclosed by the related user and/or his/her friends. For instance, let v_A , v_B , and v_C be three Facebook users and v_A discloses that he is a son of v_B , while v_C reveals that she is the wife of v_B . Using these connections, we infer that v_A is also

a son of v_C , even though v_C does not explicitly declare this relationship publicly. We would like to note that we perform this inference only for one step to avoid making biased inferences, because some users may share some loved ones as their family members even though they are not real family members. This can be explained with the following example: Assume that v_A discloses that he/she is a sibling of v_B (even this is not true in real life), and it is also inferred that v_C is a sibling of v_A . Running the inference mechanism for more than one step causes have false relationship that v_C is a sibling of v_B which is not true in real life. Therefore, the inference phase is restricted to running just one step on the overall network to eliminate this kind of false inference.

4.2. Tie Strength Calculation

A common point among existing studies on tie strength calculation over OSN data is that their models output a strength label (e.g., strong, weak) or a score representing the social strength between a pair of OSN users. A score of tie strength is computed by using weight values (i.e., parameter importance) of a set of components which may vary depending on the completeness of the OSN data at hand. On the other hand, as stated in Section 2, different techniques (e.g., linear functions, regression models) are employed to learn/compute the weights of components.

In this paper, we use real-world Facebook data that only provides whether there is a friendship link between two users without any indication of the strength of this link or tie. Since supervised learning approaches require labeled training data, we adopt and use a linear function to measure tie strength between users as in [49], [53]. Our model uses weights of the tie strength components which

Table 4.
Tie strength components

Component	Code
Bidirectional components	
Whether user v_A and user v_B have the same gender	SGEN
# of common friends	CFRI
# of liked pages in common	CPAGE
Whether user v_A and user v_B work in the same company	SCOMP
Whether user v_A and user v_B go/went to the same school	SEDU
Whether user v_A and user v_B are from the same hometown	SHT
Whether user v_A and user v_B live in the same place	SLI
Whether user v_A and user v_B have the same job title	SJOB
Unidirectional components	
# of events that user v_A attended with user v_B	EVNT
Whether user v_A reveals being in a relation with user v_B	PREL
Whether user v_A reveals being family members with user v_B	FMEM
# of posts in which user v_A tagged user v_B	PATAG
# of posts in which user v_A tagged only user v_B	PSTAG
# of directly commented posts (i.e., user v_A writes the first comment for any post by user v_B)	CMDIR
# of indirectly commented posts (i.e., user v_A writes a comment in any order but not the first for any post by user v_B)	CMIND
# of directly replied comments (i.e., user v_A writes the first reply for a comment by user v_B)	RPDIR
# of indirectly replied comments (i.e., user v_A writes a reply in any order but not the first for a comment by user v_B)	RPIND
# of direct posts shared by user v_A on the wall of user v_B	DRPST

are presented in Table 4. Inspired by [61], in this paper, we adopt and use the self-information model to assign weights to our components. This is one of the major contributions of this paper. For a set of C components $\{\lambda_1, \lambda_2, \dots, \lambda_c\}$, we construct a binary vector B_{e_q} to represent a directed edge e_q from user v_j to $v_{j'}$ (i.e., $v_j \rightarrow v_{j'}$):

$$B_{e_q} = \langle \mathcal{I}_{\lambda_1}(e_q), \mathcal{I}_{\lambda_2}(e_q), \dots, \mathcal{I}_{\lambda_c}(e_q) \rangle \quad (1)$$

where $\mathcal{I}_{\lambda_i}(e_q)$ is a component indicator function

which indicates whether e_q satisfies component λ_i :

$$\mathcal{I}_{\lambda_i}(e_q) = \begin{cases} 1 & \text{if } e_q \text{ satisfies component } \lambda_i \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Then, the tie strength between v_j and $v_{j'}$ is calculated as the weighted sum of individual component values:

$$tie_strenght(e_q) = \sum_{i=0}^c O_{\lambda_i(e_q)} \times W(\lambda_i) \quad (3)$$

where $O_{\lambda_i(e_q)}$ is the observed value of component λ_i for edge e_q . On the other hand, $W(\lambda_i)$ represents the weight of the component λ_i and equals to its self-information which is defined as $I(m) = -\log Pr(m)$ where $Pr(m)$ represents the probability that message m is chosen from all possible choices in the message space. By the way, $W(\lambda_i) = I(\mathcal{I}_{\lambda_i}(\cdot) = 1) = -\log Pr(\mathcal{I}_{\lambda_i}(\cdot) = 1)$ where it is not possible to get the true value of $Pr(\mathcal{I}_{\lambda_i}(\cdot) = 1)$. As such, we obtain the value of $W(\lambda_i)$ by estimating $Pr(\mathcal{I}_{\lambda_i}(\cdot) = 1)$ on our directed edge set E as follows [61]:

$$W(\lambda_i) = \hat{Pr}(\mathcal{I}_{\lambda_i}(\cdot) = 1) = \frac{|\{e_q \in E \mid \mathcal{I}_{\lambda_i}(e_q) = 1\}|}{|E|} \quad (4)$$

4.3. Privacy Risk Measurement

4.3.1 General Framework

Most of the existing studies measure privacy risk based on a general framework proposed by Liu and Terzi [1]. This framework depends on the sensitivity and visibility of attributes and is formulated as follows:

$$PR(v_j) = \sum_{i=1}^z PR(t_i, v_j) = \sum_{i=1}^z \beta_{t_i} \times V(t_i, v_j) \quad (5)$$

In equation 5, $PR(v_j)$ represents overall privacy risk of user v_j , while t is the total number of items considered in privacy scoring. On the other hand, $PR(t_i, v_j)$, β_{t_i} , and $V(t_i, v_j)$ correspond to privacy

risk of user v_j due to item t_i , sensitivity of item t_i , and visibility of item t_i with respect to the user v_j , respectively. As seen from the general framework, privacy scores depend on the sensitivity and visibility of items. Sensitivity measures how much sensitive information is revealed by a user, while visibility measures how wider a range of information about a user spreads. These two components are computed based on the response matrix R .

4.3.2 IRT Scoring

Liu and Terzi calculate the sensitivity and visibility of items using two different methods [1]. The first one, called the naive method, uses observed visibility and sensitivity of items and is formulated as follows in dichotomous case:

$$PR(v_j) = \sum_{i=1}^z \underbrace{\frac{n - |r_i|}{n}}_{\beta_i} \times \underbrace{\frac{|r_i|}{n}}_{P_{ij=V(i,j)}} \times \underbrace{\frac{|r^j|}{z}}_{z} \quad (6)$$

In equation 6, n and z represents the number of users and items in network respectively. On the other hand, $|r_i|$ denotes the number of users who set $r(j,i)=1$, while $|r^j|$ is the number of items for which user v_j sets $r(j,i) = 1$.

Sensitivity values computed with the naive method are significantly biased by the user population. The second method, inspired by IRT tries to remedy this problem. Risk scoring in the dichotomous case is carried out as follows [1]:

$$PR(v_j) = \sum_{i=1}^z \beta_i \times \underbrace{\frac{1}{1 + e^{\alpha_i(\theta_j - \beta_i)}}}_{P_{ij=V(i,j)}} \quad (7)$$

Notice that the equations given above are applicable in the dichotomous case. The reader can refer to [1] for details on the handling of polytomous R .

4.3.3 Intrinsic Privacy Risk Score

In their recent work, Pensa and diBlasi [31] proposed a less expensive formulation of Liu and Terzi's naive model (see eq. 6) that can be employed in both polytomous and dichotomous cases. According to their method, for any visibility degree $h = \{0, 1, \dots, l\}$, sensitivity σ_i of item t_i can be computed depending on whether h equals one of the two extreme (i.e., $h = 0$ or $h = l$) values. If h takes one of the two extreme values, the sensitivity of the item can be computed as follows [31]:

$$\sigma_{i0} = \frac{n - \sum_{j=1}^n f(r_{ji} \geq 1)}{n}, \text{ for } h = 0, \text{ and} \quad (8)$$

$$\sigma_{il} = \frac{n - \sum_{j=1}^n f(r_{ji} \geq l)}{n} \text{ for } h = l. \quad (9)$$

If the value of h is not equal to any extreme value, on the other hand, the sensitivity is computed as follows [31]:

$$\sigma_{ih} = \frac{1}{2} \left(\frac{n - \sum_{j=1}^n f(r_{ji} \geq h)}{n} + \frac{n - \sum_{j=1}^n f(r_{ji} \geq h+1)}{n} \right) \quad (10)$$

In equations given above (eq. 8, eq. 9, and eq. 10), f_A is an indicator function with value 1 when condition A is true and 0 when A is false. This two way computation of sensitivity guarantees that $\sigma_{i0} < \sigma_{i1} < \dots < \sigma_{il}$. The visibility calculation considers visibility v_{ij} (i.e., $V(i, j)$) of an item t_i due to user v_j and, for any degree $h = \{0, 1, \dots, l\}$, and it can be computed as follows [31]:

$$v_{ijh} = \frac{\sum_{j=1}^n f(r_{ji}=h)}{n} \times \frac{\sum_{i=1}^z f(r_{ji}=h)}{z} \times h \quad (11)$$

After computing sensitivity and visibility components, intrinsic privacy risk $\rho_p(v_j)$ for any given user v_j is computed as follows [31]:

$$\rho_p(v_j) = \sum_{i=1}^z \frac{\rho_p(v_j, t_i)}{\max_{v_j' \in V} \rho_p(v_j', t_i)} \quad (12)$$

In equation 12, $\rho_p(v_j, t_i) = \sum_{h=0}^l \sigma_{ih} \times v_{ijh}$ is the risk of item t_i due to user v_j .

In this paper, we use both IRT scoring and intrinsic privacy risk (IPS) scoring to compute the privacy risks of users due to their disclosed items. For the rest of this paper, we will refer to these two methods as IRT and IPS in short.

4.3.4 Network-aware Privacy Risk Score

Privacy risk of a user does not only stem from his/her publicly available information, but the risk is also affected by his/her friends [31]–[33]. For instance, if a user is mostly befriended by privacy-unaware users, then the user should be assigned a higher privacy risk than a user who is befriended by users who care about both their own and their friends' privacy [31]. As such, Pensa and diBlasi [31] proposed a network-aware privacy scoring method inspired by the page rank algorithm [62].

In this method, each user $v_j \in V$ is associated with his/her intrinsic privacy score $\rho_p(v_j)$. Unlike the original page rank algorithm, the authors use a personalized non-uniform page rank vector $P = [p(v_1), \dots, p(v_n)]^T$, where each component corresponding to node v_j is equal to $\rho_p(v_j) / \sum_{j'=1}^n \rho_p(v_{j'})$ [31]. Based on this setting, network-aware privacy scores of all users are defined as follows [31]:

$$P = dA^T P + \frac{1-d}{\sum_{j=1}^n \rho_p(v_j)} \rho \quad (13)$$

In equation 13, $\rho = [\rho_p(v_1), \dots, \rho_p(v_n)]^T$, d is a damping factor, A is an $n \times n$ matrix such that each element $a_{jj'} = a_{j'j} = 1/\text{deg}(v_j)$, if $(v_j, v_{j'}) \in E$, and $a_{jj'} = 0$ otherwise.

In this paper, we additionally, use this network aware-risk scoring method to obtain privacy risk

scores of users due to their social connections in the network. In the rest of this paper, we refer to the network-aware risk scoring method with NPS.

5. Experimental Results

To measure the privacy risk of users, we first created our response matrices using different models. In this phase, we use R1, R2, and R3 matrices populated by L1, L1 + L2, and L1 + L2 + L3 models respectively in an incremental way. Note that the L2 and L3 models infer private attributes of users based on different techniques (see Figure 1) such as regular expressions, social connections, lexicons, and meta-data of activities depending on the attribute at hand.

We performed extensive experiments on the crawled data that is described in Table 1 and contains approximately 5.9M wall activities and profile information of 20K users. The following subsections present the results of each sub-task in our flowchart, a qualitative assessment of risk scoring methods, and an overall privacy risk scoring discussion for Turkish Facebook users, respectively.

5.1. Attribute inference

Figure 2 depicts the number of revealed (for L1) and inferred (for L2 and L3) attributes for users. As seen from Figure 2, when we consider overall inferred attributes, our results show that the most inferred ones are FL, G, HT, and LIV attributes.

Using the L2 model, we were able to perform the inference task for the majority of items. The attributes whose values are most inferred are FL, G, EDU, HT, LIV, and OP, while the least inferred ones are MA and PV. In this model, KIN and PV attributes are difficult to correctly interpret through automatic tools because people are prone to use

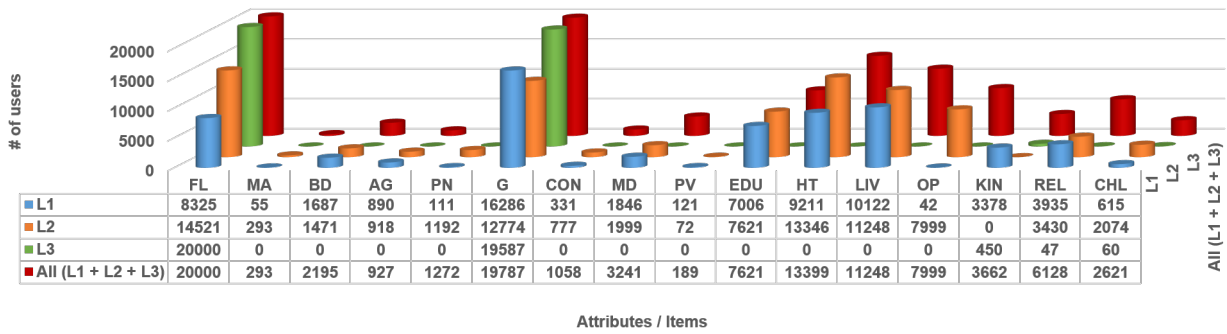


Figure 2. The number of revealed (L1) and inferred (L2 and L3) attributes with respect to the users.

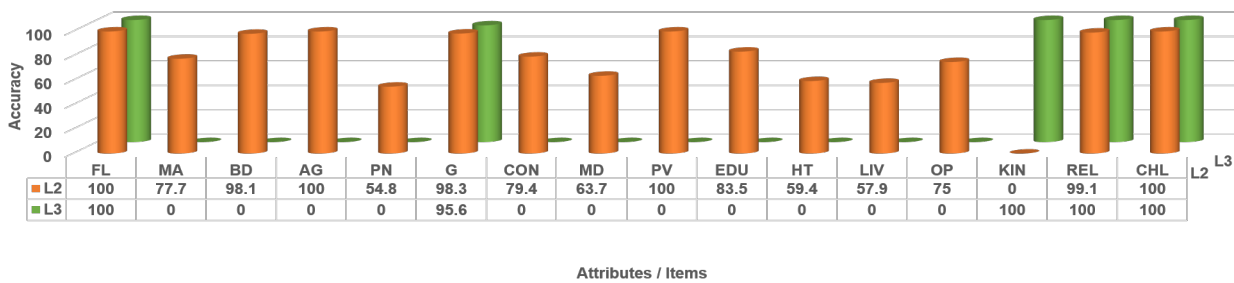


Figure 3. Accuracy of our inference models with respect to distinct attributes.

sarcastic phrases or provoking words that cannot be easily interpreted. Therefore, we prefer not to include these items in our content-based inference task in the L2 model (PV attribute is inferred just based on liked pages of users) and leave them untouched for future improvements.

The L3 model, on the other hand, can be employed to infer FL, G, PV, EDU, HT, LIV, OP, KIN, REL, and CHL attributes. However, we employed the L3 model just for a few of these attributes because users' information in our crawled snapshot is incomplete. The results in Figure 2 show that it is possible to infer the complete list of a user's friend list (i.e., FL), family membership (i.e., KIN, and CHL), and private relationship (i.e., REL). The

popularity of an attribute within the overall network can help to infer some attributes like G, even when the first neighborhood (i.e., direct friends) of a user is private or incomplete.

After completing attribute inference, we measured how accurate our L2 and L3 models are. To do so, we used a simple approach such that if there is an intersection between revealed and inferred value(s), we assume that the related model's inference is true and false otherwise. Using this approach, we measured the accuracy of our inference models concerning the attributes. The inference accuracy of an attribute is computed by dividing the number of total users whose attributes are inferred correctly by the total number of users who disclose the attribute.

Table 5.
 Computed weights of tie strength components
 in ascending order.

Component	Weight	Component	Weight
PREL	0.000577	CMDIR	0.027054
EVNT	0.001148	RPDIR	0.038379
SJOB	0.002426	SEDU	0.044330
FMEM	0.002461	CMIND	0.122535
RPIND	0.002575	CPAGE	0.194213
DRPST	0.007238	SHT	0.211385
PSTAG	0.014422	SLI	0.235391
PATAG	0.014914	CFRI	0.292170
SCOMP	0.018558	SGEN	0.769739

Inference accuracy for all of the considered attributes is given in Figure 3, where 0 means that we do not include the related attribute in our inference task for the corresponding model.

As seen from Figure 3, the L2 model often achieves considerable accuracy even though it infers attribute values from textual contents and liked pages of users. On the other hand, attributes that can not be detected using a regular format or some specific rules were inferred with lower accuracy than the accuracy of attributes inferred by regular expressions and specific rules. Attributes that are inferred with the highest accuracy by the L2 model are FL (100%), PV (100%), CHL (100%), REL (99.1%), G (98.3%), and BD (98.1%), while the lowest accuracy measurements are obtained for HT (59.4%), LIV (57.9%), and PN (54.8%) respectively.

The L3 model, on the other hand, achieves very high accuracy for all attributes (i.e., FL, G, KIN, REL, and CHL) it considers. This is because the L3 model uses social connection information that is not required to be mined or extracted by an additional step. In other words, this information may be an

edge known to be present (due to the bi-directional nature of Facebook) or an attribute revealed by any user in the network. As a final step of this first phase, we created three different response matrices in *dichotomous* form by populating revealed or inferred values of attributes. We aim to explore how the privacy risks of users were affected by our attribute inference models. For this purpose, we used an incremental way and created three response matrices, namely R1, R2, and R3 by populating them with attribute values obtained by L1, L1 + L2, and L1 + L2 + L3 respectively.

We would like to note that while populating R2 and R3 matrices, we ignored attributes that are predicted with an accuracy rate below 90%. Our aim in doing this is to filter out attributes -inferred with low accuracy- that may cause any bias in privacy risk scoring.

5.2. Results of tie strength calculation

In the second phase of our privacy scoring framework, we computed tie strengths between each pair of befriended users and injected these strengths as edge weights in the next step. We first calculated self-information of the tie strength components as their weights using Eq. 4. We obtained our weights based on the satisfaction of our components by each directed edge of 402.3K edges. Table 5 presents the results of this experiment which shows that the top five most important components of social friending are SGEN, CFRI, SLI, SHT, and CPAGE.

Secondly, we computed tie strength between each pair of users with Eq. 3 which is based on the weight and the observed value of each component in interactions between these users. We then injected these tie strengths into the graph representation of our crawled network to represent edge weights. Note that our tie strength computation produces non-symmetric weights for each (now, directional) edges

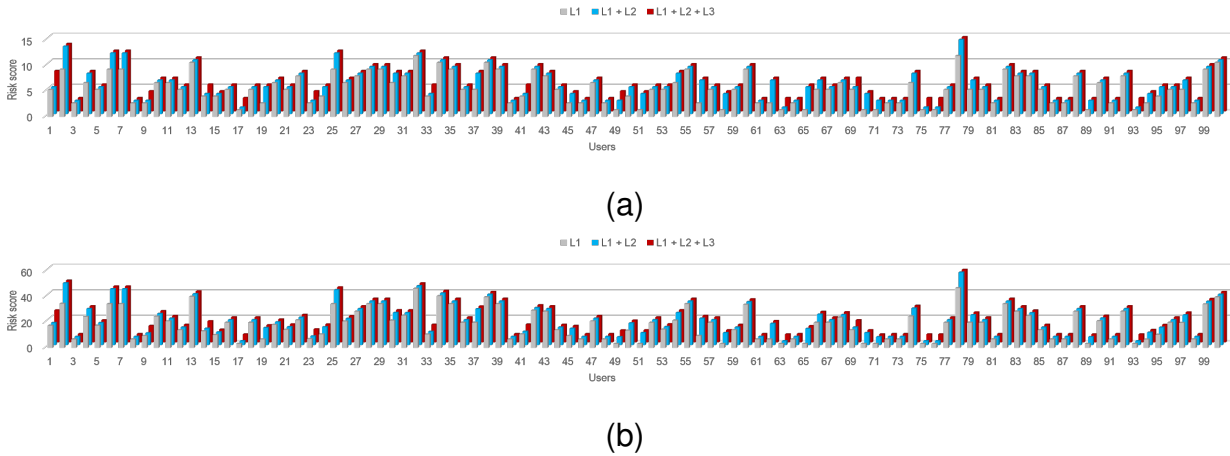


Figure 4. Privacy risks of randomly selected 100 users with respect to the L1, L1 + L2, and L1 + L2 + L3 models. Risk scores are obtained with (a) IPS, (b) IRT.

between a pair of users due to the unidirectional components given previously in Table 4.

5.3. Results of privacy risk scoring

5.3.1 Risk due to revealed and inferred attributes

In this phase, we employ IRT and IPS privacy risk scoring methods formulated in Eq. 7 and Eq. 12, respectively. We obtained privacy risks of 20K users due to their revealed and inferred items. Next, we fed risk scores into the PageRank algorithm and obtained network-aware privacy scores. Note that our response matrices (i.e., R1, R2, and R3) are populated in an incremental way and store information about how users make their items visible or think that sensitive. However, each of these matrices has different nature due to we populate them by including inferred items (actually kept private in reality) in each step. This leads an item to have different sensitivity and visibility values for each of our response matrices.

For instance, assume that an arbitrary attribute

A1 is kept private by the majority of users and we inferred its value by applying the L2 model for most of the users. In this case, the sensitivity of A1 would be high in response matrix R1 filled by the L1 model, but its sensitivity would be low in response matrix R2 filled by the L1 + L2 model. This is because item A1 would be shared in R2 by the majority of users in the OSN. This case causes to compute biased privacy risks for three models. Therefore, in this paper, we obtained sensitivity and visibility of items based on the R1 matrix and used these values to compute privacy risks from R2 and R3 matrices as well. Figure 4a and Figure 4b depict privacy risks of randomly selected 100 users due to their shared/inferred items concerning the R1, R2, and R3 matrices filled by L1, L1 + L2, and L1 + L2 + L3 models, respectively.

As seen from these figures, IRT and IPS methods have similar behavior and produce scores at different scales. Privacy risk of users increases (even though attributes inferred with low accuracy are filtered) after making attribute inference. The highest privacy risk scores are obtained on the R3 matrix which is filled by all our models.

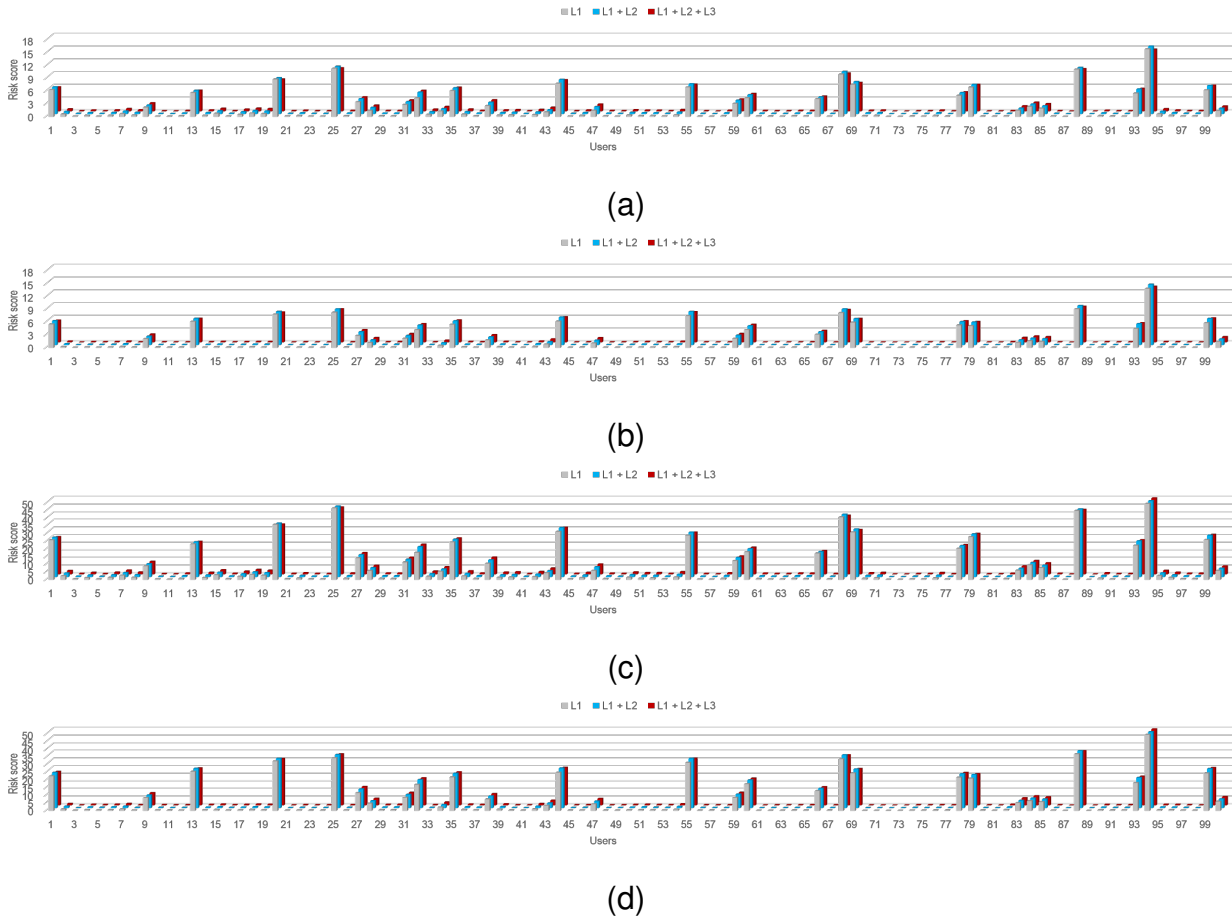


Figure 5. Privacy risks of randomly selected 100 users with respect to the L1, L1 + L2, and L1 + L2 + L3 models. Risk scores are obtained with (a) NPS_{IPS} , (b) NPS_{IPS-TS} , (c) NPS_{IRT} , and finally, (d) NPS_{IRT-TS} .

5.3.2 Risk due to friendship connections

In this phase, we obtained the NPS values of users by using the PageRank algorithm which is formulated in Eq. 13. However, differently from [31], we also injected social tie strengths (TS in short) of users as edge weights which is one of the major contributions of this paper. In this phase, we obtained NPS values in different ways where we fed IRT and IPS scores as initial values of the PageRank algorithm. As such, we use NPS_{IPS} and NPS_{IRT} to represent NPS scores obtained by using IRT and IPS scores as initial values respectively.

Additionally, we also add the “TS” term as an indicator of whether NPS scores are obtained by injecting social tie strengths or not. For instance, NPS_{IPS-TS} means that NPS scores are obtained by running PageRank injected with user’s social tie strengths and uses IPS scores as its initial values.

We obtained users’ NPS values depending on the initial risk scoring method (i.e., IRT or IPS), response matrix, and injection status of the PageRank algorithm. Figure 5a and Figure 5c depict NPS_{IPS} and NPS_{IRT} privacy risk scores of previously selected 100 users with respected to the attribute models. Figure 5b and Figure 5d, on the other hand,

show results of the same experiments on injected PageRank with social tie strengths.

As seen from these figures, privacy risks due to revealed items and social connections are significantly different from each other. Using IRT and IPS values to obtain PageRank-based NPS values only changes the scale of risk score as the behavior of the PageRank is the same for each case. NPS scores make it clear to observe which users are more central and build more social connections in the network. For instance, the user numbered 94 is at more central and builds more social connections among previously selected 100 users in the network. Tie strengths also have a significant effect on the NPS scores of users. The NPS values of users may increase or decrease depending on the tie strengths of their connections. For instance, the NPS value of the user numbered 25 decreased after injecting the PageRank with tie strengths between his/her connections, while the reverse is also true for user numbered 13 (see Figure 5b and Figure 5c).

We explore the correlation between risk scoring methods and chi-square testing in the next subsections to give a more clear insight into the behavior and the effect of our suggestion of using tie strength in NPS calculation.

5.3.3 Correlation between risk scoring methods

In this step, we computed the Pearson correlation between each pair of risk scores obtained by using these different cases. We obtained risk scores on the response matrix R1 and presented results in Table 6. As seen from Table 6, NPS values have a very low correlation with IPS and IRT values. This means that NPS values have different distributions due to network structure, and intense friendship connections have a strong effect on privacy risk

Table 6.
 Pearson correlation scores for privacy scoring methods on response matrix R1.

Method	IPS	NPS _{IPS}	NPS _{IPS-TS}	IRT	NPS _{IRT}	NPS _{IRT-TS}
IPS	1	0.21	0.24	0.98	0.22	0.25
NPS _{IPS}	0.27	1	0.79	0.20	0.99	0.79
NPS _{IPS-TS}	0.24	0.79	1	0.23	0.80	0.99
IRT	0.98	0.20	0.23	1	0.21	0.23
NPS _{IRT}	0.22	0.99	0.80	0.21	1	0.79
NPS _{IRT-TS}	0.25	0.79	0.99	0.23	0.79	1

scoring. IPS and IRT values, on the other hand, have a very high correlation which shows that these two methods have similar behavior as stated in previous subsections. Using tie strengths affects the correlation of NPS values which means that social tie strengths should not be ignored when making privacy scoring of OSN users.

5.3.4 Importance of tie strengths in NPS calculation

As seen from Table 6, network structure has a strong effect on the privacy risk scores of users. The privacy score of a user depends on his/her location in the network and the number of his/her connections with other users. However, it also depends on the weights of each connection of the users as well. As seen from Figure 5b and Figure 5d, privacy scores of users increase or decrease depending on weights (i.e., tie strength) of their connections. To explore whether there is a strong dependence between NPS score and social tie strength, we use chi-square test statistics. Our aim here is to explore whether users with high risk have a high social tie strength with others. For this purpose, we obtained the average tie strength value for each user by taking the weights of edges coming from other users to

Table 7.

The number of users with respect to the quartiles of average tie strength and NPS values.

Grade (Risk/TS)	VLTS	LTS	MTS	HTS	VHTS
R1					
VLR	2,114	1,451	386	49	0
LR	1,176	1,392	1,143	281	8
MR	605	945	1,616	281	65
HR	80	232	828	769	912
VHR	0	5	27	953	3,015
R2					
VLR	2,075	1,452	395	75	2
LR	1,162	1,363	1,150	313	13
MR	648	940	1,588	748	76
HR	90	266	834	1,899	911
VHR	0	4	33	965	2,998
R3					
VLR	2,045	1,454	414	84	3
LR	1,225	1,325	1,125	313	12
MR	610	972	1,602	742	74
HR	95	270	825	1,891	919
VHR	0	4	34	970	2,992

the current user at hand. Next, we grouped users by considering quartiles of NPS and average tie strength values into five different groups. These groups are very low (VLTS), low (LTS), medium (MTS), high (HTS), and very high (VHTS) for tie strength values; similarly risk score groups are very low (VLR), low (LR), medium (MR), high (HR), and very high (VHR) for NPS values. Notice that in the group names TS represents tie strength, and R means risk score. Afterward, we obtained the number of users for each of these groups and created a contingency table for each model as given in Table 7.

We would like to note that Table 7 shows three different contingency tables concerning the users' social tie strengths and privacy risk scores obtained

from the R1, R2, and R3 matrices respectively. We used `scipy.stats.chisquare` package to employ chi-square test statistics and obtained statistically significant p values (i.e., $p < 0.05$) with 95% confidence level for each of three contingency tables. This shows that NPS and social tie strength values are strongly dependent and there is an association between these two variables. In other words, we reject the null hypothesis that these two variables are independent of each other.

5.4. Overall Risk Analysis of Turkish Facebook users

In this section, we conduct risk analysis for all users in our crawled snapshot for investigating the privacy awareness of Turkish Facebook users and its change by gender and age. Privacy risks due to the revealed (and inferred) items and social connections have very low correlation (see Table 6) and they are at different scales (see Figure 4 and Figure 5). This is because IPS and IRT methods perform scoring just based on revealed attributes by users. Network aware scoring method (i.e., NPS), on the other hand, considers network structure.

As a result, these two approaches produce different privacy scores for each user. As such, in this analysis, we take the average privacy risk scores obtained with these two approaches. Even though IPS and IRT methods produce highly correlated scores (see Table 6), we selected IRT as our intrinsic risk scoring method. This is because IRT approximates the item characteristics curve that best represents the response matrix [1]. To obtain network-aware privacy risk scores we fed IRT scores as initial values of the PageRank both with tie strengths (NPS_{IRT-TS}) and without tie strengths (NPS_{IRT}) to make the effect of tie strengths more clear.

To make a reasonable analysis, we first scaled averaged risk scores in a range of 0 and 1. Then

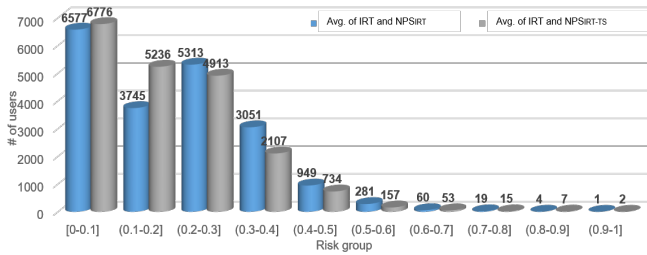


Figure 6. The number of users in ten different groups with respect to average of IRT and NPS values

grouped users into ten different groups based on their risk scores. We would like to note that for this analysis we obtained IRT scores on the response matrix R3. We present the results of this experiment in Figure 6 which shows that majority of users have less than 0.5 risks of privacy and using social tie strengths has a strong effect on the privacy risk of users. To exemplify, the number of users having risk scores in a range between 0.1 to 0.2 is 3,745 without tie strengths, whereas it is 5,236 when tie strengths are considered in NPS calculation.

Next, we explored the relative percent of 11,830 male and 7,957 female users (in total 19,787 users whose gender attributes were disclosed or inferred, see Figure 2) concerning their gender attributes and average privacy risk scores. Note that in the rest of our experiments, we use the average of IRT and NPS_{IRT-TS} values. We show the results of this experiment in Figure 7, which depicts that male users often have a higher risk of privacy than females. Note that in this experiment, we considered users who reveal their gender attributes along with users whose gender attributes are inferred by the L2 and/or L3 model(s).

Finally, we investigated the relative percent of 206 users younger than 30 years old and 721 users

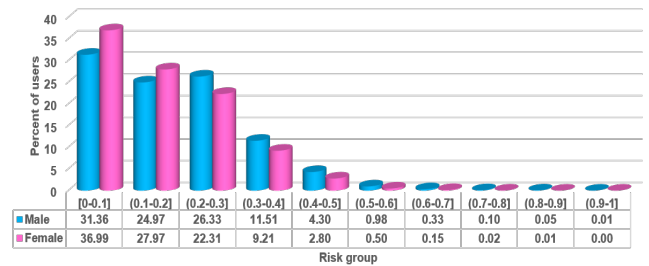


Figure 7. Relative percents of users in ten risk groups with respect to average privacy risk and gender attribute.

who are 30 years of age and older with respect to their age and average privacy risk attributes. In this experiment, we again considered users whose age is obtained from the revealed birth date (i.e., BD) attribute and others whose age is inferred by the L2 model.

As seen from Figure 8, in a range of the privacy risk between 0 and 0.4, the relative percent of users who are under 30 years of age is greater than the relative percent of those users over the 30 years of age. On the other hand, in a range of the privacy risk between 0.4 and 1.0, relative percentages of users over 30 years of age are greater than those of users under 30 years of age. These results show that users over thirty years of age often have a lower risk than other users. However, the majority of users in high-risky groups (i.e., involving risk scores of 0.4 and higher) are over the age of thirty. This is highly because users over 30 years of age and older both share more sensitive information and build more social connections compared to younger users.

We would like to note that we do not name our risk groups in the experiments in this subsection due to group names (e.g., high, low, very high, etc.) are relative and can vary from person to person. However, our privacy analysis of Turkish Facebook

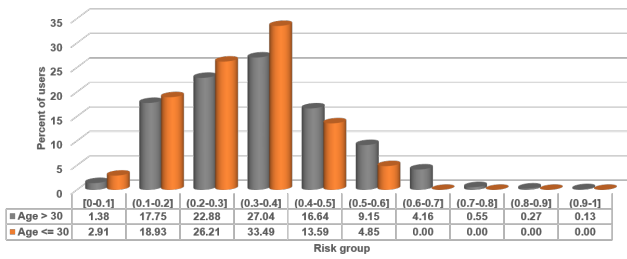


Figure 8. Relative percents of users in ten risk groups with respect to average privacy risk and age attribute.

users shows that majority of them have less than 0.5 risks of privacy. On the other hand, male users have a higher privacy risk than females, and users over the age of 30 years take place in high-risky groups more frequently than younger users in general.

6. Discussion

In this paper, we perform a privacy risk analysis of Turkish Facebook users employing a novel risk scoring framework that relies on attribute inference to produce an R matrix as input. It then uses both intrinsic risk scores along with social tie strengths to produce its final output representing network-aware risk scores of users. For this purpose, we use real-world OSN data unlike most of the existing studies. In addition, we perform attribute inference to fill our dichotomous response matrices.

Our attribute inference results show that many attributes of OSN users can be learned by using their textual contents and social connections. Our results related to attribute inference show that the majority of Turkish Facebook users keep their attributes private, but they are still not aware that an attribute can be learned or inferred by using different techniques. Therefore, they often disclose their sensitive information by posting activities on their walls or

building connection with other users. However, the risk is so serious that a user may contribute to putting any other user at risk even though he/she is not befriended with him/her by only disclosing his gender. This is because an adversary can infer a user’s gender by using the popularity of his/her first name among other male and female users in the network.

Users disclose their private attributes in their activity contents or other text-based fields of their profiles. Even though we used a restricted list of attributes in our inference tasks, textual contents of users have a great potential to infer various other attributes (e.g., health status, hobbies, etc.) of users. Our results in the content-based inference model (i.e., the L2 model) provided low accuracy for some of our attributes. However, taking the accuracy metric as the only value to measure the performance of an inference task may not be sufficient. This is because the inferred value of an attribute may be true even though it does not match disclosed value. For instance, a user may have many phone numbers and only disclose one of them in the related field of his profile. However, this user can post an activity including his/her other phone numbers at the same time. In such a case, using only the accuracy measure can mislead us because the inferred value may actually be true.

Inference based on user connections (i.e., the L3 model) provides very accurate results and it is easy to perform compared to the content-based model. Besides, in this model the more complete the OSN data, the higher the inference performance. It will also be possible to apply inference tasks for more attributes in more complete OSN data.

OSN users have privacy risk scores due to their revealed items and social connections in the network. In the literature, centrality or network-aware risk scoring methods ignore social tie strengths between

users. In this paper, we inject and use social tie strengths in network-aware risk computation. We also show that there is a strong association between high-risk scores and social tie strength. Therefore, we suggest not just using the social connections of users in network-aware risk scoring, but also using weights (i.e., tie strengths) of the social connections of users. Please note that computationally there is not an overload (taking apart the attribute inference and assuming that the R matrix is populated) within our framework that employs existing studies in a different manner. Its disadvantage may be producing risk scores with a two-step calculation, but the advantage is that, unlike existing studies, our framework considers social tie strengths. We believe that this way of scoring produces more concrete and reliable risk scores for users.

7. Conclusion

In this paper, we perform privacy scoring of Turkish Facebook users by proposing a framework that involves our novel aspects together with existing methods. These novel aspects are populating the R matrix by using attribute inference and obtaining network aware-risk scores not just by using users' connections but weights of these connections as well. In the attribute inference phase, we use two different ways of inference mechanisms (i.e., L2, and L3) that completely rely on textual content and social connections of users. In this phase, we use several dedicated regular expressions, rules, and activity meta-data information to infer the attributes of users. Afterward, we use both disclosed (L1) and inferred attributes (L2 for inferred from contents and L3 for inferred from connections) to derive three different response matrices (i.e., R1, R2, and R3) in an incremental way in which the response matrices are populated by L1, L1 + L2, and L1 + L2 + L3 models respectively. We would like to

note that inferring attributes both from contents and connections is one of the first major contributions of this study in the context of Turkish OSN users. Upon completion of the attribute inference, we discarded some of the attributes within the both L2 and L3 models having an inference accuracy lower than 0.9 (i.e., 90%) to prevent having biased results. Using these sets of features, we first obtained intrinsic privacy risk scores that were then used to compute network-aware risk scores of users. In the phase of network-aware risk scoring, employing the self-information model to compute weights of tie strength components is the second major contribution of this paper.

Based on our results, we conclude that the social tie strengths of users' connections have to be taken into consideration when computing their centrality or network-aware risk scores. What is more, it is often possible to infer the private attributes of users by using network structure information. However, network structure is not enough to perform inference in some cases due to OSN data being often incomplete. The content-based inference is, on the other hand, very challenging compared to inference based on network structure. This is because OSN textual data is often dirty and generally written in informal language. Performing such an inference task without a learning model requires many rules, sources, and so on. Using a learning model, on the other hand, requires a labeled corpus for each related inference task. When a learning model is employed, it may improve the accuracy of attribute inference for some attributes like gender; however, inference of some other attributes may still be challenging like EDU due to inferring such items require combining and applying multiple learning tasks such as named entity recognition, word sense disambiguation, and machine learning.

In future work, we are planning to perform a sen-

sitivity classification of users' wall activity contents to detect whether a given textual content is sensitive or not in terms of privacy. Additionally, we will try to propose a hybrid method that handles the PRE by combining the powers of existing risk scoring methods and classical text categorization. One of the possible scenarios to exemplify is employing term weighting approaches to compute sensitivity or visibility calculation of attributes.

Acknowledgements

The authors would like to thank the anonymous reviewers for their useful comments and suggestions.

Appendix

A summary of the studies with a focus on privacy risk scoring over OSNs is given in Table 8, where the column *data* indicates the source of OSN data used in experiments. We clearly specify whether the data is synthetic or real, and additionally mark survey-based input data with (*S*) and textual content data with (*C*). *Type of information* column indicates what kinds of OSN data are utilized in a study. Here, *profile items* refer to basic profile information; *actions* refer to user interactions in the form of posting data, liking someone's posts, making new connections and sharing; and *content* refers to the actual (typically textual) content of a user's share. *POV* column stands for point-of-view of the corresponding study. We mark studies with a combination of letters *U*, *O*, *F* and *S* to respectively imply that the study is centered on user, OSN, friend(s) and strangers (i.e., users who are not friends of account owner). Finally, the *PRE* column indicates whether or not the study utilizes a privacy risk estimation method - a novel method of its own or an existing method.

References

- [1] K. Liu and E. Terzi, "A framework for computing the privacy scores of users in online social networks," in *Ninth IEEE International Conference on Data Mining*, December 2009, pp. 288–297.
- [2] A. Srivastava and G. Geethakumari, "Measuring privacy leaks in online social networks," in *IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, August 2013, pp. 2095–2100.
- [3] M. Sramka, "Evaluating privacy risks in social networks from the user's perspective," in *Advanced Research in Data Privacy. Studies in Computational Intelligence*, G. Navarro-Arribas and V. Torra, Eds. Cham: Springer, 2015, pp. 251–267.
- [4] C. Akcora, B. Carminati, and E. Ferrari, "Privacy in social networks: How risky is your social graph?" in *IEEE 28th International Conference on Data Engineering*, April 2012, pp. 9–19.
- [5] S. Oukemeni, H. Rifà-Pous, and J. M. M. Puig, "Ipam: Information privacy assessment metric in microblogging online social networks," *IEEE Access*, vol. 7, pp. 1–20, 2019.
- [6] E. Aghasian, S. Garg, and J. Montgomery, "A privacy-enhanced friending approach for users on multiple online social networks," *Computers*, vol. 7, no. 3, pp. 1–12, 2018.
- [7] J. Caramujo and A. M. R. da Silva, "Analyzing privacy policies based on a privacy-aware profile: The facebook and linkedin case studies," in *IEEE 17th Conference on Business Informatics*, July 2015, pp. 77–84.
- [8] Y. Yang, J. Lutes, F. Li, B. Luo, and P. Liu, "Stalking online: on user privacy in social networks," in *Proceedings of the second ACM conference on Data and Application Security and Privacy*, February 2012, pp. 37–48.
- [9] J. Alemany, E. del Val, J. Alberola, and A. García-Fornes, "Estimation of privacy risk through centrality metrics," *Future Generation Computer Systems*, vol. 82, pp. 63–76, 2018.
- [10] E. Aghasian, S. Garg, L. Gao, S. Yu, and J. Montgomery, "Scoring users' privacy disclosure across multiple online social networks," *IEEE access*, vol. 5, pp. 13 118–13 130, 2017.
- [11] N. Talukder, M. Ouzzani, A. K. Elmagarmid, H. Elmeleegy, and M. Yakout, "Privometer: Privacy protection in social networks," in *IEEE 26th International Conference on Data Engineering Workshops (ICDEW)*, March 2010, pp. 266–269.
- [12] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior," *Proceedings of the National Academy of Sciences*, vol. 110, no. 15, pp. 5802–5805, 2013.
- [13] A. Mislove, B. Viswanath, P. K. Gummadi, and P. Druschel, "You are who you know: inferring user profiles in online social networks," in *Proceedings of Web Search and Data Mining*, February 2010, p. 251–260.
- [14] B. S. Vidyalakshmi, R. K. Wong, M. Ghanavati, and C. H. Chi, "Privacy as a service in social network communications," in

- IEEE International Conference on Services Computing*, February 2014, pp. 456–463.
- [15] M. H. Veiga and C. Eickhoff, “Privacy leakage through innocent content sharing in online social networks,” arXiv preprint arXiv:1607.02714, 2016.
- [16] A. Srivastava and G. Geethakumari, “Privacy landscape in online social networks,” *International Journal of Trust Management in Computing and Communications*, vol. 3, no. 1, pp. 19–39, 2019.
- [17] B. S. Vidyakshmi, R. K. Wong, and C. H. Chi, “Privacy scoring of social network users as a service,” in *IEEE International Conference on Services Computing*, July 2015, pp. 218–225.
- [18] L. Bioglio and R. G. Pensa, “Impact of neighbors on the privacy of individuals in online social networks,” *Procedia Computer Science*, vol. 108, pp. 28–37, 2017.
- [19] P. Gundecha, G. Barbier, and H. Liu, “Exploiting vulnerability to secure user privacy on a social networking site,” in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, August 2011, pp. 511–519.
- [20] T. Minkus and N. Memon, “On a scale from 1 to 10, how private are you? scoring facebook privacy settings,” in *Proceedings of the Workshop on Usable Security*, February 2014, pp. 1–6.
- [21] A. Srivastava and G. Geethakumari, “A privacy settings recommender system for online social networks,” in *International Conference on Recent Advances and Innovations in Engineering*, May 2014, pp. 1–6.
- [22] Q. Wang, H. Xue, F. Li, D. Lee, and B. Luo, “# donttweetthis: Scoring private information in social networks,” *Proceedings on Privacy Enhancing Technologies*, vol. 2019, no. 4, pp. 72–92, 2019.
- [23] J. A. Biega, K. P. Gummadi, I. Mele, D. Milchevski, C. Tryfonopoulos, and G. Weikum, “R-susceptibility: An ir-centric approach to assessing privacy risks for users in online communities,” in *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, July 2016, pp. 365–374.
- [24] X. Song, X. Wang, L. Nie, X. He, Z. Chen, and W. Liu, “A personal privacy preserving framework: I let you know who can see what,” in *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*, July 2018, pp. 295–304.
- [25] A. C. Islam, J. Walsh, and R. Greenstadt, “Privacy detective: Detecting private information and collective privacy behavior in a large social network,” in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, November 2014, pp. 35–46.
- [26] X. Li, Y. Yang, Y. Chen, and X. Niu, “A privacy measurement framework for multiple online social networks against social identity linkage,” *Applied Sciences*, vol. 8, no. 10, pp. 1–19, 2018.
- [27] H. Mao, X. Shuai, and A. Kapadia, “Loose tweets: an analysis of privacy leaks on twitter,” in *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society*, October 2011, pp. 1–12.
- [28] K. Thomas, C. Grier, and D. M. Nicol, “unfriendly: Multi-party privacy risks in social networks,” in *International Symposium on Privacy Enhancing Technologies Symposium*, July 2010, pp. 236–252.
- [29] O. Coban, A. Inan, and S. A. Ozel, “Privacy risk analysis for facebook users,” in *28th Signal Processing and Communications Applications Conference*, October 2020, pp. 1–4.
- [30] R. G. Pensa and G. D. Blasi, “A privacy self-assessment framework for online social networks,” *Expert Systems with Applications*, vol. 86, pp. 18–31, 2017.
- [31] R. G. Pensa, G. D. Blasi, and L. Bioglio, “Network-aware privacy risk estimation in online social networks,” *Social Network Analysis and Mining*, vol. 9, no. 1, pp. 1–15, 2019.
- [32] R. G. Pensa and G. D. Blasi, “A semi-supervised approach to measuring user privacy in online social networks,” in *International Conference on Discovery Science*, October 2016, pp. 392–407.
- [33] —, “A centrality-based measure of user privacy in online social networks,” in *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, October 2010, pp. 255–265.
- [34] J. Domingo-Ferrer, “Rational privacy disclosure in social networks,” in *International Conference on Modeling Decisions for Artificial Intelligence*, October 2010, pp. 255–265.
- [35] E. M. Maximilien, T. Grandison, T. Sun, D. Richardson, S. Guo, and K. Liu, “Privacy-as-a-service: Models, algorithms, and results on the facebook platform,” in *Proceedings of W2SP 2009: Web 2.0 Security and Pivacy 2009*, May 2009, pp. 1–4.
- [36] I. Symeonidis, F. Beato, P. Tsormpatzoudi, and B. Preneel, “Collateral damage of facebook apps: an enhanced privacy scoring model,” IACR Cryptology ePrint Archive, 2015.
- [37] O. Coban, A. Inan, and S. A. Ozel, “Inverse document frequency-based sensitivity scoring for privacy analysis,” *Signal, Image and Video Processing*, vol. 16, pp. 735–743, 2022.
- [38] A. Braunstein, L. Granka, and J. Staddon, “Indirect content privacy surveys: measuring privacy without asking about it,” in *Proceedings of the Seventh Symposium on Usable Privacy and Security*, July 2011, pp. 1–14.
- [39] A. Djoudi and G. Pujolle, “Social privacy score through vulnerability contagion process,” in *Fifth Conference on Mobile and Secure Services (MobiSecServ)*, March 2019, pp. 2–3.
- [40] J. L. Becker, “Measuring privacy risk in online social networks,” M.S. thesis, Dept. of Comp. Sci., University of California, Sacramento, CA, USA, 2009. [Online]. Available: <https://www.proquest.com/docview/304853196?pq-origsite=gscholar&fromopenview=true>
- [41] Y. Zeng, Y. Sun, L. Xing, and V. Vokkarane, “A study of online social network privacy via the tape framework,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 7, pp. 1270–1284, 2015.

- [42] H. Simo, H. Shulman, M. Schufirin, S. L. Reynolds, and J. Kohlhammer, "Privinervis: Towards enhancing transparency over attribute inference in online social networks," in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops*, May 2021, pp. 1–2.
- [43] L. Bioglio, S. Capecchi, F. Peiretti, D. Sayed, A. Torasso, and R. G. Pensa, "A social network simulation game to raise awareness of privacy among school children," *IEEE Transactions on Learning Technologies*, vol. 12, no. 4, pp. 456–469, 2019.
- [44] A. Halimi and E. Ayday, "Real-time privacy risk quantification in online social networks," *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 74–81, 2021.
- [45] O. Coban, A. Inan, and S. A. Ozel, "Towards the design and implementation of an osn crawler: a case of turkish facebook users," *International Journal of Information Security Science*, vol. 9, no. 2, pp. 76–93, 2020.
- [46] O. Coban, "An exploratory analysis of leaked facebook data: A case of turkish users," *International Journal of Information Security Science*, vol. 10, no. 4, pp. 119–137, 2021.
- [47] C. G. Akcora, B. Carminati, and E. Ferrari, "Risks of friendships on social networks," in *IEEE 12th International Conference on Data Mining*, December 2012, pp. 810–815.
- [48] I. Kahanda and J. Neville, "Using transactional information to predict link strength in online social networks," in *Third International AAAI Conference on Weblogs and Social Media*, May 2009, pp. 74–81.
- [49] S. Krakan, L. Humski, and Z. Skocir, "Determination of friendship intensity between online social network users based on their interaction," *Tehnicki vjesnik*, vol. 25, no. 3, pp. 655–662, 2018.
- [50] Z. Liu, H. Li, and C. Wang, "New: A generic learning model for tie strength prediction in networks," *Neurocomputing*, vol. 406, pp. 282–292, 2020.
- [51] Y. D. Seo, Y. G. Kim, E. Lee, and D. K. Baik, "Personalized recommender system based on friendship strength in social network services," *Expert Systems with Applications*, vol. 69, pp. 135–148, 2017.
- [52] S. S. Rodriguez, R. P. D. Redondo, A. F. Vilas, and J. J. P. Arias, "Using facebook activity to infer social ties," in *Proceedings of the 2nd International Conference on Cloud Computing and Services Science*, April 2012, pp. 325–333.
- [53] J. Ilic, L. Humski, D. Pintar, M. Vranic, and Z. Skocir, "Proof of concept for comparison and classification of online social network friends based on tie strength calculation model," in *6th international conference on information society and technology*, March 2016, pp. 159–164.
- [54] Y. Kılıç and A. I. A., "Qpr-eval: A quantitative framework for privacy risk score evaluation," unpublished.
- [55] O. Coban, "Attribute inference over real-world online social networks: a comprehensive privacy analysis," Ph.D. dissertation, Dept. of Comp. Sci., Cukurova Univ., Adana, Turkey, 2021.
- [56] O. Coban, A. Inan, and S. A. Özel, "Facebook tells me your gender: An exploratory study of gender prediction for turkish facebook users," *Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 4, pp. 1–38, 2021.
- [57] A. Canaydın, "Github Repo: alpcanaydin/liseler," Accessed May. 12, 2021. [Online]. Available: <https://github.com/alpcanaydin/liseler/blob/master/liseler-web/public/data.json>
- [58] Anonymous, "Github Repo: Liseler-Ilkokullar," Accessed May. 13, 2021. [Online]. Available: <https://github.com/SqlHareketi/Liseler-Ilkokullar/blob/master/liselerilkokullar.sql>
- [59] F. Sarhan, "Archived Github Repo: ButunUniversiteListesiCrawler," Accessed Jun. 12, 2021. [Online]. Available: <https://github.com/f9n/ButunUniversiteListesiCrawler/blob/master/universities.txt>
- [60] D. Irani, M. Balduzzi, D. Balzarotti, E. Kirda, and C. Pu, "Reverse social engineering attacks in online social networks," in *International conference on detection of intrusions and malware, and vulnerability assessment*, July 2011, pp. 55–74.
- [61] Y. Wang, T. Liu, Q. Tan, J. Shi, and L. Guo, "Identifying users across different sites using usernames," *Procedia Computer Science*, vol. 80, pp. 376–385, 2016.
- [62] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer networks and ISDN systems*, vol. 30, no. 1-7, pp. 107–117, 1998.

Table 8.

An overview of studies devoted to evaluate and measure the privacy in OSNs. The POV and PRE stand for the point of view and privacy risk estimation respectively.

Study	Data	Type of Information			Method / Framework	POV	PRE	Year
		Profile Items	Actions	Content				
[1]	Synthetic & Survey Data	✓			IRT and Naive Models	U	✓	2009
[2]	Item Response Matrix (S)	✓		✓	Privacy Quotient based on [1]	U	✓	2013
[3]	Real & Multiple OSNs Data	✓		✓	Extended Concept of [1]	U	✓	2015
[4]	Real Facebook Data	✓	✓	✓	Sight	S	✓	2012
[5]	Real & Multiple OSNs Data				IPAM	O	✓	2019
[6]	Synthetic Facebook & Twitter Data	✓			Privacy-Enhanced Friending	U	✓	2018
[7]	Real OSN Data	✓			Privacy-aware UML profile	O	✗	2010
[8]	Real LinkedIn & Phonebook Data	✓			Two Attacker Models	U	✗	2012
[9]	Synthetic & Real Twitter (PHEME)		✓		PRS	U	✓	2018
[10]	Real & Multiple OSNs Data	✓			PDS	U	✓	2017
[11]	Real Facebook Data	✓			Privometer	F	✓	2010
[14]	Real Facebook Data			✓	A Privacy Service Model	O	✗	2014
[15]	Real & Multiple OSNs Data			✓	Informativeness Score	U	✗	2016
[17]	Synthetic Facebook Data	✓	✓	✓	PS based on FACT	F	✓	2015
[18]	Synthetic Facebook Network		✓		Extension of SIR model	U	✗	2017
[20]	Surveyed Facebook Data (S)		✓		Extension of method from [35]	U	✓	2014
[19]	Real Facebook Data	✓	✓		I-Index	U	✓	2011
[22]	Real Twitter Data (C)		✓	✓	Privscore	U	✓	2019
[23]	Real & Multiple Communities Data (C)			✓	R-Susceptibility	U	✓	2016
[24]	Real Twitter Data (C)			✓	TOKEN	U	✓	2018
[25]	Real Twitter Data (C)			✓	Privay Detective	U	✓	2014
[26]	Survey, Real & Multiple OSNs Data	✓		✓	Privacy score based on [1]	U	✓	2018
[27]	Real Twitter Data (C)			✓	Based on Machine Learning	✗	✗	2011
[28]	Real Facebook Data (C)		✓	✓	Based on Machine Learning	✗	✗	2010
[30]	Real Facebook Data (S)	✓	✓		Extension of method from [1]	U	✓	2017
[31]	Simulated & Real Facebook Data (S)	✓	✓		Extension of method from [1]	U	✓	2019
[34]	Simulated Data	✓			PFS based on [1]	U	✓	2010
[35]	Survey Data (S)	✓			PaaS (Naive model from [1])	O	✓	2009
[36]	Real Facebook Data	✓			PET based on [1]	U	✓	2015
[37]	Real & Synthetic Data	✓			IDF-based scoring	U	✓	2022
[38]	Survey Data (S)	✓		✓	Statistical Methods	U	✗	2011
[39]	Enron Email Data		✓		Contagion Process Model	U	✓	2019
[21]	Real Facebook Data	✓		✓	PrPu	F	✓	2014
[40]	Real Facebook Data	✓			PrivAware	U	✓	2009
[41]	Real Facebook Datasets	✓	✓	✓	TAPE	U	✓	2015
[42]	Real Data	✓	✓		PrivInferVis	U	✓	2021
[43]	Predefined Set of Posts			✓	Social4school	U		2018
Ours	Real Facebook Data	✓	✓	✓	Extension of network-aware risk scoring from [31] (see Figure 1)	U	✓	2022