Privacy in Online Social Networks

Elie Raad, Richard Chbeir

Privacy in Online Social Networks

Elie Raad[1] and Richard Chbeir[2]

[1] Memorial University of Newfoundland, Canada
elie.raad@mun.ca,
[2] University of Pau and Adour Countries, Laboratoire LIUPPA, France
richard.chbeir@univ-pau.fr

Abstract. Online social networks have become an important part of the online activities on the web and one of the most influencing media. Unconstrained by physical spaces, online social networks offer to web users new interesting means to communicate, interact, and socialize. While these networks make frequent data sharing and inter-user communications instantly possible, privacy-related issues are their obvious much discussed immediate consequences. Although the notion of privacy may take different forms, the ultimate challenge is how to prevent privacy invasion when much personal information is available. In this context, we address privacy-related issues by resorting to social network analysis and link mining techniques. We first describe the fundamental of social networks, their common representations, and the main motivations associated with their use. Afterwards, we particularly show how privacy attacks can build on social network analysis and link mining techniques to reveal user-sensitive information. The chapter concludes with a discussion of some open challenges to address in future privacy-related works.

Keywords: Social Networks, Privacy, Social Network Analysis, Link Mining, Knowledge Discovery, Social Network Representation;

1 Introduction

For the past few years, online social networks experienced an exponential growth in the number of their users and in the huge amount of available information. Many online social networks like Facebook[3], LinkedIn[4], Google+[5], and Twitter[6] offer to web users new interesting means to communicate and interact. In reality, information available on these networks commonly describes persons along with their personal information (e.g., what they like, where do they live, who they know, etc.) and interactions (e.g., with who they exchange messages, what comments they post, how often they update their personal status, etc.).

[3] http://www.facebook.com
[4] http://www.linkedin.com
[5] https://plus.google.com/
[6] http://www.twitter.com

With the proliferation of online social networks, information sharing on these networks is gaining an ever-increasing importance. Obviously, online social networks have found ingenious ways to collect data as users socialize. Not surprisingly, when socializing users communicate, interact, and tend to freely reveal personal information in line with their perceptions and preferences. To control the access to this personal information and to enforce its protection, online social networks promote the use of a number of built-in control mechanisms [1] [2]. However, social network users often fail to fully protect their profiles and personal data from undesirable forms of access as it has been revealed by previous studies [3] [4] [5] [6]. This is due to the limited efficiency of the provided control mechanisms [1] [6] [7] and to the users' misconceptions about the networks' composition, the visibility of their profiles, and their misunderstandings of the privacy risks [3] [8].

As a result, more and more accessible personal information is available online, and yet, though the risk of security breaches and data exposures are manifold, adequate tools and efficient solutions are still missing. Social network users, overwhelmed with information, struggle to properly maintain privacy over their data to meet their actual expectations. However, social network users are not security experts and do not fully control their data. It is even hard to handle privacy threats as privacy breaches become numerous when dealing with personal information posted over years, across many online social networks, and shared with different types of contacts (e.g., colleagues, relatives, friends, etc.). In addition, existing privacy settings are relatively complicated to be correctly managed by users [6] [5], online social networks may suffer from design conflicts issues (security and privacy vs. usability and sociability) [9], and may intentionally or accidentally leak users' information to unauthorized entities or third parties [10]. Consequently, it is vital to protect the tremendous amount of information from all sorts of attacks that may compromise users' privacy, invade their security, or disclose their data to unauthorized parties. Therefore, providing effortless mechanisms for social network users allowing them to control and reduce the potential exposure of their private information is of valuable importance.

With the huge number of social network users, it is therefore complex to delineate the concept of privacy. Privacy is a topic that received a lot of attention and has different facets [9] [11] [12] [13]. However, on online social networks some key characteristics that underly privacy are commonly identified [13]. Among these concepts *anonymity*, *unlinkability*, and *unobservability* are the most interesting. Firstly, *anonymity* ensures that an attacker cannot sufficiently identify a user within a set of users. Secondly, *unlinkability* refers to the incapacity of an attacker to distinguish whether two or more items of interest are related or not. Thirdly, *unobservability* protects a user's activity so that an attacker or a third party cannot tell whether a resource or service is being used. Today, those are the most common users' privacy concerns.

In this chapter, we discuss privacy on social networks which is one of the most intriguing social networks' challenges. We argue that in order for a system to provide optimal privacy, its underlying algorithms must understand the char-

acteristics of social networks and their associated analysis techniques. We also focus on the importance of social network data and explain how network analysis and data mining techniques [14] [15], useful in understanding users' behaviors and networks' characteristics, can become a source of privacy risk. On social networks, privacy concerns seem to be world-wide challenges for users, and thus novel privacy protection techniques must provide clear answers to a multitude of questions surrounding privacy:

- What are the most adequate analysis tools to use when dealing with specific privacy concerns?
- To which extent social networks' users behaviors are comparable to real-world persons' behavior and how to avoid that?
- How to prevent unwanted information leakage, data exploitation, and information linkage?
- What are the most important elements in order to protect users' privacy (e.g., type of data exchanged, relationship types, networks structures, etc.)?

This chapter is structured as follows. We begin by defining the fundamental concepts of social networks in Section 2. We then focus on online social networks in Section 3 and we describe the main components of these networks as well as the main motivations associated with their use. This is followed by a description of the appropriate ways to represent social networks in Section 4 where we particularly highlight the graph-based representation. In Section 5, we discuss the challenges and opportunities related to the availability of social network data, its protection, its analysis, and list some privacy protection techniques. We then present social network analysis in Section 6, a particularly important research area to study networks. We describe its most commonly used measures and their associated privacy threats. In Section 7, we detail link mining and its different tasks that are also used to analyze networks while emphasizing on social network links. We highlight the characteristics of the various link mining tasks and show their derived privacy threats. In Section 8, we present some open challenges yet to be addressed in future privacy-related systems before concluding this chapter in Section 9.

2 What is a Social Network?

Networks have been used to model many systems of interest such as the World Wide Web [16], computer networks [17], biochemical networks [18], diffusion networks [19], and social networks [20]. Each of these networks is a structure that consists of a set of actors representing, for instance, web pages on the World Wide Web or persons in a social network, connected together by relations, representing links between web pages or friendships between persons. Besides these structural properties (actors and relations), Wasserman and Faust [14] identified a number of fundamental concepts like ties, dyads, triads, subgroups, and groups, that characterize networks. For the purpose of this work, we start by detailing the concepts of actors, relations, and ties, the building blocks of social networks, before illustrating their use in online social networks (Section 3).

Definition 1 *An actor is a social entity that interacts with other entities not only to maintain existing relations but also to establish new ones. On social networks, the concept of actors can refer to various types of entities such as persons, groups, and organizations.*

Actors interact with each other through a variety of meaningful relations that denote different patterns of communication. Relations like friendship, collaboration, and alliance can vary across time, applications, or in terms of the involved actors [21]. Consequently, there are two main categories of networks that can be identified based on the type of actors, one-mode networks and two-mode networks [22]. While one-mode networks have a single type of actors, two-mode networks, also called bipartite, are networks with two types of actors. For instance, social networks modeling friendship between actors are an example of one-mode networks whereas those concerned with group memberships or attendance at events are two-mode networks.

Definition 2 *A relation represents a connection from one actor to another one. A relation, also called relationship, plays an important role when studying the structure of social networks and the interactions among their actors. A relationship is characterized by various features such as its content, direction, and strength.*

The relationship types have been addressed in several studies. Borgatti et al. [23] distinguished between four basic types of relationships: similarities, social relations, interactions, and flows. For instance, these relationships can express memberships (e.g., same club), kinships (e.g., mother of), affections (e.g., likes), interactions (e.g., talked to), and flows (e.g., flow of information), among others. Relationships on social networks can be directed or undirected. Depending on their content, relationships may (or may not) have a specific direction. While relationships such as "marriage" and "friendship" are undirected, other relationships such as "parent of" or "fan of" are directed. Social network relationships can also differ in strength. Usually, the strength can be estimated in a variety of ways using information about the actors, their interaction activities, or the correlation between them as the most common indicators [21] [24] [25].

Definition 3 *A tie is the set of all relationships that exist between two actors. It is tightly connected to the concept of relationship as it aggregates the different types of relationships that exist between two actors. Just like relationships, ties also vary in terms of their content, direction, and strength.*

Actors can be connected either with one relationship exclusively (e.g., employees of the same company) or with many relationships (e.g., employees of the same company and members of a sport club at the same time). Consequently, pairs of actors who maintain more than a single relationship are said to have a tie [26] [27]. While each individual relationship within a tie carries its own content and direction, the strength of a tie depends on many factors such as

the number of relationships that actors maintain, the reciprocity of these relationships, and their duration. Granovetter [28] distinguished between strong and weak ties on the basis of the time actors spend together, their intimacy, and the emotional intensity of the existing relationships. Generally, weak ties are infrequently maintained with little interactions among actors (e.g., between distant acquaintances). Strong ties link similar actors, such as close friends, whose social circles tightly overlap with each other. Often, actors with strong ties that maintain many kinds of relations tend to communicate frequently with each other and use different channels of communication [29].

The previously defined concepts (actors, relations, and ties) are particularly important to understand and to study social networks. Besides the fact that social networks are made of several components, online social networks can also hold different types of data and can have various representations as detailed in the next sections.

3 Online Social Networks

Interactions between actors and offline communications between persons have always been central in the study of social networks [30] [31] [32]. Many studies investigated ties between friends and relatives in order to understand why actors provide different types of social support [30], how social networks are formed, persist and disappear [31], and what methods are appropriate in order to estimate the size of personal communication networks [32]. More recently, the impact of social-based technologies on users, and particularly the influence of online social networks, is becoming the major source of contemporary fascination and controversy [27] [33] [34]. A number of studies shed the light on different research directions like the implications of online social networks on individual connectivity [35], the capacity of technology to override cognitive limits in order to socialize with larger groups [36], and the challenge to maintain a balance between security, privacy, usability, and sociability on online social networks [9] [12]. Our focus in this chapter is on the privacy aspects of social networks, where research has primarily aimed to protect social network users with their profiles and relationships. In the following, we first highlight the main motivations associated with social networks' use and show some relevant statistics. We then describe the concepts of social network users, user profiles, and social relationships specifically in the context of online social networks. Note that in the following, we refer to social networks and online social networks interchangeably.

3.1 Motivations and Use

Social networks and content-sharing sites with social networking functionalities have become an important part of the online activities on the web and one of the most influencing media. Facebook, LinkedIn, Twitter, MySpace[7], Flickr[8], and

[7] http://www.myspace.com/
[8] http://www.flickr.com/

Youtube[9] are among the most popular online social networks. These networks are attracting an ever-increasing number of users, many of whom are interested in establishing new connections, maintaining existing relations, and using the various social networks' services. Facebook, for instance, reported to have one billion monthly active users[10] that are uploading more than 250 million photos every day. On Twitter, 8 terabytes of data is generated on Twitter per day[11]. Another study published by Nielsen[12] on social networking reported that social networks and blogs dominate the time that users spend on the web and now account for nearly 20% of the total time spent online on personal computers and 30% of online time on mobile devices such as smartphones and tablets. With the huge number of users and the tremendous amount of shared data, such social networks will indisputably shape the future of online communication.

A large and growing body of literature has investigated the influence of using social networks on users' interactions [37] [38], gratifications [39] [40], self-estime [41] [42], and sharing practices [43] [44]. Recent studies examined the use of social networks, the behaviors that surround online interactions, and the benefits perceived by social network users [45] [46] [47] [48]. The findings of these studies show that the motivations for using social networks are numerous. They indicate that the enjoyment is the most influential factor [49], followed by the users' interest to frequently interact with their real-world life friends [50], and the founded users' belief that social networks improve the efficiency of their shared information to enforce existing connections and to connect with new users [51].

3.2 Social Network Users

While many definitions exist for the term social network [52] [53] [54], all of them are centered around social network users. First, these users create a personal profile which usually contains identifying information (e.g., name, age, photos, etc.) and captures users' interests (e.g., joining groups, liking brands, etc.). Afterwards, users start to socialize by interacting with other network members using a wide variety of communication tools offered by different social networks. In reality, each social network offers particular services and functionalities to target a well-defined community in the real world. Many of these available services are designed to help foster information sharing [55], bridge online and offline connections to enforce interactions [56], provide instant information help [46], and enable users to derive a variety of uses and gratifications from these sites [39]. To make use of the provided functionalities and to stay tuned with their related members, users create several accounts on various social networks where they disclose personal information with varying degrees of sensitivity [57]. Personal information available on these networks commonly describes users and their interactions, along with their published data.

[9] http://www.youtube.com/

[10] https://www.facebook.com/press/info.php?statistics, accessed 01 October 2012

[11] http://www.information-management.com/issues/21_5/big-data-is-scaling-bi-and-analytics-10021093-1.html, accessed 01 October 2012

[12] http://blog.nielsen.com/nielsenwire/social/2012/

3.3 User Profiles

Information about each social network user is maintained in a user profile which contains a number of attributes related to the demographics of users, their personal and professional addresses, their interests and preferences, as well as different types of user-generated contents (e.g., posts, photos, videos, etc.) [58] [59]. Prior studies have noted the importance of user profiles to shape users' personalities, identities, and behaviors on social networks [7] [41] [42]. These studies showed that among the disclosed attributes such as personal information and user-generated contents, photos and status updates have higher preferences for users. User profiles also store the contact lists that consist of various interpersonal relationships as discussed in the following. Currently, social network sites do not all adopt the same user profile attributes' representation. Different technologies provide users with an extensive list of attributes to describe their profiles such as:

- **RDFa**[13]**:** standing for *Resource Description Framework - in - attributes*, is a W3C recommendation used to embed semantic into XHTML. RDFa is a thin layer of markup that can be added to web pages and make them more understandable for machines as well as for persons. RDFa provides a consistent syntax and big expressivity by proposing an integration of the RDF triple concept (subject, predicate, attribute) with the flexible XHTML language, which is used by web browsers.
- **Microformats**[14]**:** are little pieces of structured information embedded into XHTML documents. They transform documents to machine-readable semantic data such as contact details, social relationships, event information, etc. Currently, different microformats exist for different needs such as hCard used to describe persons, companies, and organizations with a limited set of elements representing business cards, calendars for events (e.g., hCalendar), decentralized tagging (e.g., rel-tag), etc.
- **XFN**[15]**:** standing for *XHTML Friends Network*, represents 18 human relationships with a set of values and gives the possibility to authors, for example, to indicate which of the weblogs they read belong to friends they have met.
- **FOAF**[16]**:** standing for *Friend Of A Friend*, is a machine-readable semantic vocabulary describing persons, their relationships, and activities. FOAF documents are written in XML syntax and adopt the conventions of the Resource Description Framework (RDF)[17]. Among the many representations, FOAF is considered as the richest vocabulary to use in terms of describing users' profiles and has currently become a widely accepted standard [60]. FOAF defines a set of attributes, grouped into categories as shown in Figure 1. A sample FOAF profile is illustrated in Figure 2.

[13] http://www.w3.org/MarkUp/2009/rdfa-for-html-authors
[14] http://microformats.org
[15] http://gmpg.org/xfn
[16] http://xmlns.com/foaf/spec
[17] http://www.w3.org/RDF

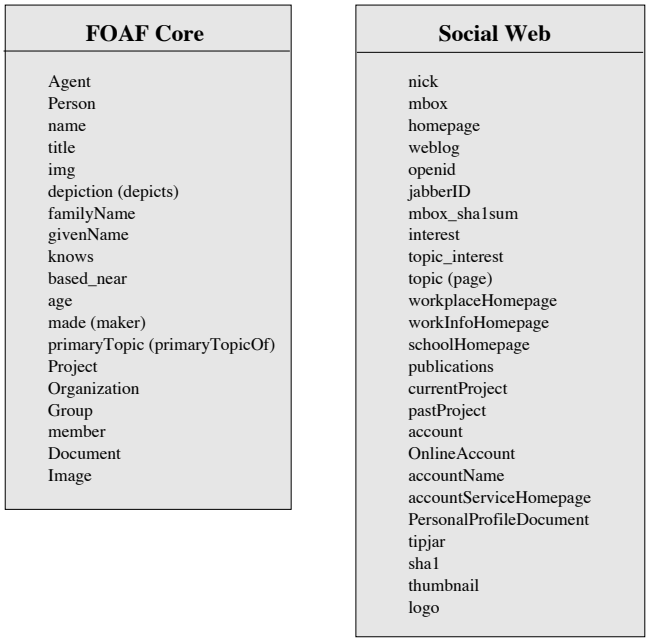


Fig. 1. Main FOAF attributes grouped into categories: FOAF Core and Social Web.

3.4 Social Relationships

While myriad social networks' services assist users to find new contacts and establish new connections (e.g., friend suggestion systems through locations [61], based on interactions [21], etc.), users get connected to different types of contacts such as friends, relatives, colleagues, and strangers. Nevertheless, social relationship types between users and their contacts are rarely identified neither by the users nor by the existing social network sites [62] [63] [64]. This diversity, yet the different levels of social closeness between users and their contacts, entails an increasing need to analyze social interactions for better relationship (and consequently privacy) management. Currently, users are often provided with an exclusive and default relationship type connecting them to each of their contacts within a single social network site. However, it is common that social network users initiate connections with other contacts without any prior offline connection [65]. On Facebook, for instance, these contacts are known as *friends* even though social network users do not particularly know or trust them. Consequently, many privacy-related concerns are raised in terms of identity disclosure, information sharing, access control, etc. [9]. The default social relationship(s) among the users of a number of famous social networks, along with other information, can be found in Table 1. Given the diverse sources of social relationships, further research is needed to better understand the privacy needs of users as more friendships continue to be forged and maintained with

```
<foaf:Person>
<foaf:name>Alexandre William</foaf:name>
<foaf:firstname>Alexandre</foaf:firstname>
<foaf:family_name>William</foaf:family_name>
<foaf:mbox rdf:resource=aw@somesite.com/>
<foaf:homepage rdf:resource=http://personalsite.com/aw/>
<foaf:workplaceHomepage rdf:resource=http://workaddress.com/>
<foaf:img>www.xyz.com/alex/photos/alex.jpg</foaf:img>
<foaf:interest>Paris, Software, Internet</foaf:interest>
<foaf:knows><foaf:Person>
<foaf:mbox rdf:resource=contact1@somesite.com />
<rdfs:seeAlso rdf:resource=http://contact1.net/foaf.rdf./>
</foaf:Person></foaf:knows>
<foaf:knows><foaf:Person>
<foaf:mbox rdf:resource=contact2@somesite.com />
</foaf:Person></foaf:knows>
</foaf:Person>
```

Fig. 2. Sample FOAF document.

online and offline contacts. The structure of the networks, the user-generated content, the level of interaction, as well as other dimensions, can also be used to analyze users' behaviors and understand their privacy needs. Next, we address in detail the structural representation of social networks.

4 How to Represent a Social Network?

Finding an appropriate representation that can facilitate efficient and accurate interpretation of network data is an important step in social network studies. Just as graphs are a set of interconnected nodes, social networks are built on the foundation of actors interconnected through relationships. The use of graphs is a powerful visual tool and a formal means to represent social networks as detailed in this section.

4.1 Why Graphs?

There are many notations to represent social networks: algebraic notations, matrices, and graphs. A sample algebraic notation, a matrix representation, and a graph are illustrated in Figure 4.2. Depending on the data to be processed, the notation whose representation best fits the social network to describe is typically selected. But, there are well-known limits to the extent to which social networks can be formalized using matrices or algebraic notations to be recalled here. First, social networks hold valued relations and user-related attributes that algebraic notations cannot handle. Second, matrices are mostly efficient for small networks.

Table 1. Famous social networks with their main focus, default relationship(s), and the relationship's direction

Social Network	**Focus**	**Default Relationship(s)**	**Relationship Direction**
Facebook	General Use	Friendship	Symmetrical
Flickr	Photo-Sharing	Contact and optionally Friend or Family	Symmetrical
Google+	General Use	Friends, Family, Acquaintances and Following	Symmetrical
LinkedIn	Professional	Business	Symmetrical
MySpace	General Use	Friendship	Symmetrical
Twitter	Microblogging	Follower-Followee	Asymmetrical
Youtube	Video-Sharing	Subscribed-to	Asymmetrical

Consequently, due to the large size of social networks, matrices are not the most appropriate way to represent these networks. Note that to represent a social network using matrices, a two-way matrix, also called sociomatrix, can be used. A sociomatrix consists of rows and columns that denote social actors, and numbers or symbols in cells that denote existing relationships. Thus, graph-based representations are by far the most common form for modeling social networks [14] [66] [67]. Graphically representing social networks facilitates the understanding, labeling, and modeling of many properties of these networks (e.g., friendships networks with labeled actors and relationships). Hence, graphs can represent various social data properties and their attributes while handling large real-world networks. Beside an adequate vocabulary to denote structural properties, graph-based representations have shown their mathematical reliability as well as their capacity to prove theorems for different social structural properties [14]. More details about the advantages and drawbacks of each representation are provided in Table 2.

4.2 Graph Representation

Graphs are usually used to represent networks in different fields such as biology, sociology, and computer science [68]. Graphs consist of nodes to represent actors, and edges to represent relationships. The terms *nodes* and *objects* are usually used to denote *actors*. Likewise, *edges* may also be called *links*, or *relationships*. Nodes with multiple edges are used to represent *ties* related pairs of actors with more than one relationship.

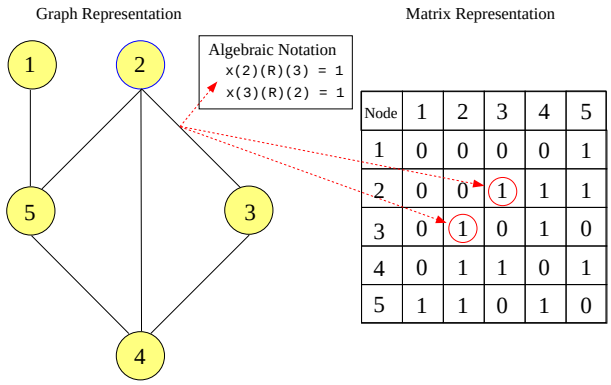
More formally, a graph, $G = (V, E)$, consists of a set of nodes, V, and a set of edges, E. The number of elements in V and E are respectively denoted as $n = \|V\|$, the number of nodes, and $m = \|E\|$, the number of edges. The *ith* node, v_i, is usually referred to by its order i in the set V. Note that E consists

Table 2. Social network representations: advantages and drawbacks

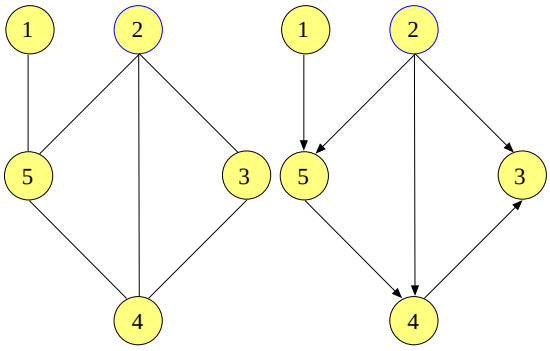
Representation	Advantages	Drawbacks
Algebraic notations	- Useful for multi-relational networks as they can easily denote the combination of relations	- Cannot handle valued relations and user-related attributes
Matrices	- Efficient for small networks - Easy to denotes ties between a set of actors (a matrix for each relationship)	- Not a best choice for large social networks - Difficult to use when network data contain information on attributes
Graphs	- Handle large social networks - Provide a rich vocabulary to easily model social networks (labels, values, weights, etc.) - Provide mathematical operations that can be used to quantify structural properties and prove graph-based theorems	- Scalable visualization techniques are needed - Signed and valued graphs have to be used to represent valued relations

of a finite set of relationships that is built from all relationships R_i, R_{i+1}, .., R_k, where k is the total number of relationships linking the pairs of actors. A subgraph *G' = (V', E')* of *G = (V, E)* is a graph such that $V' \subseteq V$ and $E' \subseteq E$. To represent different forms of data and to model the structural properties of social networks, graphs can have their edges and nodes labeled or unlabeled, directed or undirected, weighted or unweighted as explained in what follows.

Directed and Undirected Graphs In an undirected graph, the order of the connected vertices of an edge is not important. We refer to each link by a couple of nodes i and j such as *e(i, j)* or e_{ij}, i and j are the end-nodes of the link. A directed graph is defined by a set of nodes and a set of directed edges. The order of the two nodes is important: e_{ij} denotes the link from i to j, and $e_{ij} \neq e_{ji}$. To graphically indicate the direction of the links, directed edges are depicted by arrows. Depending on the nature of the relationship (asymmetric or symmetric), social network graphs can be undirected or directed. In fact, social networks can be modeled as undirected graphs when relationships between actors are mutual (e.g., symmetric relationships on Facebook where e_{ij} or e_{ji} both denote a *friendship* link between user i and user j). Social networks can also be modeled as directed graphs when relationships are not bidirectional (e.g., asymmetric relationships on Twitter where e_{ij} stands for user i is *following* user j).

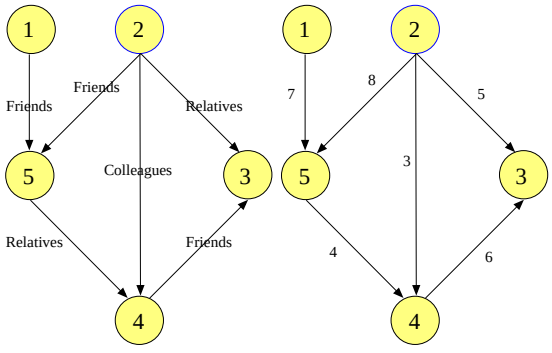


(a) Graph, Matrix and Algebraic notation sample case to represent social network actors and their relationships



(b) Undirected Graph

(c) Directed Graph



(d) Labeled Directed Graph

(e) Weighted Directed Graph

Fig. 3. A social network representation using a graph, its related matrix, and a sample algebraic notation (a), an undirected graph (b), a directed graph (c), a labeled graph (d), and a weighted graph (e) with $n = 5$ nodes and $m = 6$ links.

Figure 4.2 and Figure 4.2 show respectively a representation of an undirected and a directed graph, both with $n = 5$ and $m = 6$. Directed links are important to evaluate the role of actors in a social network. They are key factors in measuring the centrality of actors in a social network. An interesting research work conducted by Brams et al. [69] described how to transform undirected graphs to directed ones in order to explore additional information about the networks' structure. This transformation is an important step in understanding the flow of influence in the context of terrorist networks. In another study, Morselli et al. [70] investigated and compared the structure of criminal and terrorist networks. The authors used links to compute a number of measures such as degree, betweenness, and centrality measures. These measures are used in order to discover the organizational hierarchy and to identify central and powerful criminal and terrorist actors. We detail these measures in Section 6.

Labeled and Unlabeled Graphs Labels are important since they can identify the type of relationships between social network actors. When graphs are labeled, this means that a label is used to indicate the type of link that characterizes the relationship between the connected labeled nodes. Note that labeled graphs are considered to be signed graphs whenever their edges are labeled with either a + or a -. For example, a signed graph can be used to model the inferred trust or distrust relationships in online social networks [71]. Figure 4.2 shows a labeled graph where the relationship type between linked actors is indicated. On social networks, relationship can be used to organize contacts based on their relationship types. This is useful in different situations such as improving face clustering and annotation of personal photo collections [72], organizing friends into social circles [62] [64], and enforcing access control [73]. Relationship-based access control is highly interesting in order to enable users to manage and fine-tune their privacy settings.

Weighted and Unweighted Graphs Weights represent the strength of relationships between social network actors. When graphs are weighted, this means that their edges are assigned with a numerical weight, w, that can provide various indications such as link capacity, link strength, level of interaction, or similarity between the connected nodes (e.g., the number of messages that actors have exchanged, the number of common friends, etc.). Figure 4.2 shows a weighted graph (on a scale of 0 to 10) where the numeric values are assigned to the links and indicate the level of interaction between social network's actors. One way to characterize relationships is by computing their strength. On social networks, link strength is highly correlated with the level of interaction between users. Link strength can be used to model different levels of friendship where high weights represent "close friends" and low weights represent "acquaintances". Xiang et al. [74] estimated the link strength from interaction activities (e.g., communication, tagging, etc.) and user similarities. Stutzman et al. [75] argued that link strength can be used to reduce the burden of manually specifying privacy settings for each contact within a user's social network. They proposed an automated grouping of

users based on many criteria where link strength is highlighted as one of the most commonly considered factors. More recently, another research explored a more specific aspect related to the predictive capacity of link strength to generalize from one social network to another [25]. Typically, link strength is primarily used to build intelligent systems that can favor interactions with strong ties without missing interesting activities derived by weak ties. Specifically, this interesting study showed that the link strength model captured in one social network can be generalized to another network, one in which it did not train.

To sum up, structural characteristics of a graph are a key aspect for social networks as they can be used to analyze the activity and to understand the behaviors of social network users. In most cases, networks of interconnected users are mainly represented by graphs, while graphs resulting from users' activity are usually referred to as the activity graphs. The activity captured within social networks is between users (the nodes) sharing various directed or undirected relationships (the links) and different levels of interactions (strong and weak ties). In this regard, these characteristics can be used to identify well-connected, central, and influential users. This would give more visibility and understanding for the network analyzer but at the same time this can possibly reveal additional and sensitive information about the users, thus raising privacy concerns. In the next section, we discuss a number of challenges and opportunities related to the use of social networks from a user perspective.

5 Social Network Data: Opportunities and Challenges

Social networks have become an important platform for connecting users, sharing information, and a valuable source of social network data. Thus, the availability of such data represents an opportunity for people to study and analyze these networks. However, the various sources of data on social networks are not only perceived as sets of values and repositories of knowledge; rather their availability becomes a form of threat as they can be exploited by attackers to disclose various sensitive information (e.g., identities, attributes, locations, etc.). In this section, we focus on social network data and address the challenges related to how data is collected, what data is collected, how data is protected, and we list a number of existing techniques used to protect the privacy of social networks users.

5.1 How Data is Collected?

Traditionally, most of social network data were collected through questionnaires in order to study networks. These studies conducted face-to-face interviews [30], telephone surveys [76], or computer-based questionnaires [31]. To construct social networks using questionnaires, participants may spend a burdensome and unrealistic amount of time and effort in answering questions that can be difficult or repetitive. In addition, during the questionnaires participants may forget some relevant information or misinterpret questions. Consequently, such conventional methods have many limitations from different points of view related to scalability, subjectivity, inconsistency, error handling issues, etc. [77].

Today, the picture has changed. The use of electronic data extraction methods has been beneficial in collecting relevant network data, and their success spread to various domains such as hyperlink networks on the web [78], biochemical networks [18], and email messages archives [79]. In order to study social networks, novel techniques have been developed as well as adapted measures so to collect relevant data collections [80] [81] [24] [82] . Marin et al. [82] highlighted the importance of the type of networks and the type of relationships in the process of collecting network data. In their study, the authors considered two important dimensions along which network data vary: whole vs. egocentric networks, and one-mode vs. two-mode networks. Note that the difference between one-mode networks and two-mode networks is explained in Section 2. As for the difference between whole and egocentric networks, it can be simply explained by noting that the egocentric networks privilege the study of one focal node (the ego) rather than considering all the nodes of the network as in the whole network analysis [14]. Many social network systems have been developed to collect, built, and analyze data from the web such as the Referral Web [80], Flink [81], and Polyphonet [24] or from online social networks such as Twitter [83], Flickr [84], and Facebook [85]. Currently, social networks allow users to exchange various types of information, including messages, photos, and comments. Many studies have shown that social network users are highly motivated to interact with their contacts and to share personal information [43] [37] [38] [41] [42] [44]. As a matter of fact, social networks provide new possibilities to collect data more efficiently and cost-effectively.

5.2 What Data is Collected?

There are many types of social network data that can be collected from various sources on the web (i.e., different social network sites) and extracted from the daily activities and interactions between users. In this context, Schneier [86] proposed a taxonomy of social data that we further develop into two main categories:

1. **Explicit data:** is the set of explicit information that is provided by social network users or the data that is embedded in the provided information, i.e., metadata embedded in photos. Explicit information may include different forms of data such as text messages, photos, or videos. In this category, social network users actively participate in the creation of information.
   (a) **Service data:** is the set of data that a user provides to the social network to create her account such as the user's name, date of birth, country, etc.
   (b) **Disclosed data:** is what the user posts on her social network profile. This might include comments, posted photos, posted entries, captions, shared links, etc.
   (c) **Entrusted data:** is what the user posts on other users' profiles. This might include comments, captions, shared links, etc.
   (d) **Incidental data:** is what other social network users post about the user. It might include posted photos, comments, notes, etc.

2. **Implicit data:** is the set of information that is not explicitly provided by social network users. However, social networks or third parties can use the set of explicit data to infer more information about the user. Inferring implicit data is founded on the analysis of the users' behaviors or derived from one or more user-provided information. For instance, it is possible to predict the characteristics of relationships between a number of users by examining the different aspects related to the patterns of communication between users (e.g., text messages, published photos, number of common friends, etc.) [87] [62]. Consequently, in this category social network users are considered to be passive since the inferred information is extracted from prior activities or previously posted data.
   (a) **Behavioral data:** is the data inferred from the user's behaviors. Social networks can collect information about the user's habits by tracking the patterns of activities of the user and consequently analyzing the user's behavior. Inferred behavioral data can reveal various information such as what the user usually do on the social networks, with whom the user usually interacts, and in what news topics the user is interested. Social networks collect such information by analyzing the articles that the user reads, the posts that the user publishes, the game that the user plays on social networks, etc.
   (b) **Derived data:** is the data about the user that can be inferred from all other data. It is not related to the habit of the user. For example, the IP address can be used to infer the users' actual location. The derived data can also be inferred from the combination of two (or more) information. For example, if a significant number of contacts live in one city, one can say that the social network user might live there as well. In this case, social networks or third parties must have access to two information in order to infer the derived data (the contacts of a user as the first information and their corresponding hometown as the second information).

5.3 How Data is Protected on Social Networks?

Privacy on social networks is a complex concept which involves major challenges [2] [9]. A recent research in [1] addressed the topic of privacy settings on social networks and particularly investigated the privacy settings of Facebook. The results of this study show that privacy settings matched users' expectations in only 37% of the time and up to 39% when the users modified the default privacy settings [1]. The authors concluded that there is a big disparity between the desired and actual privacy settings and calls for new tools to manage privacy. This conclusion is in line with another study where social network users considered that the privacy settings are effective to manage threats coming from outside their social circles [6]. However, the same users experienced increasing concerns when it comes to sharing content with members of their social circle. Similar results were also identified when investigating privacy concerns and mechanisms surrounding tagged photos on social networks [7]. In the case of photos, the

central point behind the users' priorities associated privacy concerns with identity and impression management. In most of the time, social network users were worried of seeing an unwanted photo of them online or being tagged on an unflattering photo. All of these studies indicate that the existing privacy systems as well as their designs must be improved to better address threats and meet users' expectations.

5.4 Are Existing Privacy Protection Techniques Useful?

Privacy is closely related to network anonymization [88] [89] [90], privacy preservation [91] [92] [93] [94], and access control [73] [95] [96]. Nowadays, these conventional techniques have several disadvantages when it comes to protecting privacy of social network users.

Definition 4 *Network anonymization consists to manipulate the network's information in order to make the process of nodes' identification difficult for attackers.*

This can be achieved by modifying and removing all the attributes of the nodes in the network. However, anonymizing a network is often not enough to protect privacy since many attacks can apply de-anonymization techniques to re-identify nodes and their hidden attributes [88] [89] [90]. Attackers can use another release of the network to process the anonymized network in order to re-identify the protected nodes and consequently reveal sensitive information. As stated earlier, various sources of information are available on the web (e.g., users with accounts on different social networks, personal blogs, etc.) and thus many sources can be used as a background knowledge by the attackers. The structure of the social network and the background knowledge can be exploited by the attackers and consequently expose users to privacy issues. This is usually possible by using a variety of information such as the total number of contacts, the number of common contacts, and the relationship strength.

Definition 5 *Privacy preservation focuses on protecting sensitive information primarily through using techniques such as hiding sensitive attributes, hiding users' identities, modifying data, and randomizing values.*

Several efforts have been extensively investigating the protection of sensitive information using privacy preservation techniques [91] [92]. Besides the fact that privacy preservation has been most successful in dealing with relational data [93], privacy on social network is a confluence of several factors that when combined can lead to infer the original value of the sensitive information. All users' activities, relationships, and shared content can be potentially monitored, recorded, and analyzed by attackers. Consequently, a hidden attribute on a user's profile can still be inferred accurately. For instance, it is possible to predict the home address of a user by analyzing the geographical place of the most frequent updates posted at night or on the weekend [97]. Another study derived the user's

location given the known location of the user's friends [98]. Similarly, the availability of metadata embedded within shared content (e.g., GPS location, date, time, and device name embedded in photos) as well as the use of location-based services (e.g., Foursquare[18], Facebook, etc.) can significantly raise privacy concerns and complicate the task of protection [99]. Moreover, failing to provide an optimal identity protection can lead to disclose other sensitive information such as the type of relationships among users [94].

Definition 6 *Access control mechanisms seek to secure the access to sensitive information without explicit authorization by implementing appropriate access control mechanisms.*

On social networks, there is a growing interest in implementing access control systems but very little work has been done in these directions. Notably, social network users strive to reduce the inefficiency of current privacy systems and look forward to enforce their privacy protection. Several access control systems have been proposed, including:

- Attribute-based access control [100]: the systems in this category grant or deny access based on the user's attributes. An attribute or a set of attributes form the digital credential and may contain attributes such as age, citizenship, employment, group membership, or credit status.
- Multimedia-based access control [101]: the multimedia-based systems tend to integrate multimedia objects in the decision process. Multimedia objects can yield valuable information sensed from multimedia devices about the users and their context (e.g., user's surrounding, moves, gestures, people nearby, etc.).
- Purpose-based access control [102]: the purpose-based systems grant access to certain data with conditions. The notion of condition determines the access purpose in a dynamic manner. Such access control systems integrate the purpose of the access in the decision process and consequently dynamically associate the purpose with the requested data objects.
- Relationship-based access control [73] [95] [96]: the approaches in this category are designed to enforce users' privacy and enable users to tune their privacy settings by controlling access based on the type of relationship. For instance, the access to a specific content is authorized only for the user's colleagues, family members, etc. However, these existing works assume that relationship labels are provided with social networks. Consequently, relationship-based access control can be rarely implemented since relationship types are often missing [63] [62] [64].

5.5 Discussion

In the context of social networks, there are many challenges to overcome such as networks design and architecture, active user population and network dynamics,

[18] http://foursquare.com/

user interactions, user behavior, and most importantly privacy issues [103]. To improve users' experience while protecting their personal information is a challenge that requires adequate methods capable to analyse the different types of data, to explore the different components of social networks, and to understand the social interactions between users. As discussed in this section, the risk of disclosing private, personal and potentially sensitive information is serious since social network users lack of appropriate means to efficiently control and easily protect their published data.

In the following, we show how attackers can resort to various techniques in order to reveal sensitive information. We mainly investigate privacy concerns that are derived from social network analysis [14] and link mining techniques [15]. Our classification inspiration draws from the literature of both techniques and illustrates the privacy threats associated to common social network analysis measures and link mining tasks. We start by introducing social network analysis and link mining before detailing their corresponding privacy threats in Section 6 and Section 7, respectively.

6 Social Network Analysis: Measures and Threats

Over the past years, there has been a surge of interest in social network analysis, with works ranging from exploiting networks' structures to examining actors' roles and their interaction patterns [14]. Understanding the characteristics of social networks, namely information related to structure, is of considerable importance to deal with privacy issues, and hence social network analysis presently attracting widespread interests. In this section, we briefly trace the history of social network analysis, provide an overview of its most common measures, and discuss where and how social network analysis has been used in the context of social networks' privacy.

6.1 Development and Measures

Social network analysis has a well-established tradition in psychology and in social sciences [66]. Since the beginning of the 20th century, several social network analysis studies - primarily in educational and developmental psychology - have been conducted to study characteristics of groups (e.g., structure, formation, behavior, etc.) and social ties (e.g., influence, interaction, companionship, etc.) [104] [105] [106]. Drawing inspiration from these previous works, modern social network analysis emerged as interdisciplinary field with contributions from various areas of study such as sociology, anthropology, and mathematics [107] [108]. Concerned with the structural analysis of social interactions, modern social network analysis developed new models to study the fundamental properties of diverse theoretical and real-world networks [109]. The small-world model [110] and the scale-free model [111], useful in describing very large networks, are among the most important models that emerged from these efforts.

Social network analysis has been used in different application domains such as email communication networks [87], learning networks [29], epidemiology networks [112], terrorist networks [70], and online social networks [113]. These works tried to answer a handful of questions such as how highly an actor is connected within a network? Who are the most influential actors in a network? How central is an actor within a network? To capture the importance of actors within a network, a number of measures have been proposed in the literature [114]. A commonly accepted measure is the centrality measure.

Definition 7 *Centrality consists of giving an importance order to the actors of a graph by using their connectivity within the network.*

Several structure-based metrics have been proposed to compute the centrality of an actor within a network, such as degree, closeness, and betweenness centrality [115]. Table 3 summarizes the characteristics of these structure-based centrality measures.

Table 3. Main centrality measures and their characteristics

Centrality Measure	Characteristic
Degree	Measures how much an actor is highly connected to other actors within a network
Closeness	Computes the length of paths from an actor to other actors in the network
Betweenness	Measures the extent to which an actor lies on the paths between other actors

In what follows, we explain each of these metrics in details:

– **Degree Centrality:** Measures how much an actor is highly connected to other actors within a network. Degree centrality is a local measure since its value is computed by considering the number of links of an actor to other actors directly adjacent to it. A high degree centrality denotes the importance of an actor and gives an indication about potentially influential actors in the network. With a high degree of centrality, actors in social networks serve as hubs and as major channels of information in a network. Degree centrality, C_D, of an actor, v_i, can be computed as follows [115]:

$$C_D(v_i) = \sum_{i=1}^{n} a(v_i, v_j) \tag{1}$$

where n is the total number of actors in the social network, $a(v_i, v_j) = 1$ if and only if v_i and an actor, v_j, are connected by an edge; otherwise $a(v_i, v_j) = 0$.

– **Closeness Centrality:** Computes the length of paths from an actor to other actors in the network. By measuring how close an actor is to all other actors, closeness centrality is also known as the median problem or the service facility location problem. Actors with small length path are considered more important in the network than those with high length path. Closeness centrality, C_C, of an actor, v_i, can be computed as follows [115]:

$$C_C(v_i) = \frac{n-1}{\sum_{i=1}^{n} d(v_i, v_j)} \tag{2}$$

where n is the total number of actors in the social network, $d(v_i, v_j)$ is the geodesic distance from actor v_i to another actor v_j.

– **Betweenness Centrality:** Measures the extent to which an actor lies on the paths between other actors. It denotes the number of times an actor needs to pass via a given actor to reach another one, and thus represents the probability that an actor is involved into any communication between two other actors. Actors with high betweenness centrality facilitate the flow of information as they form critical bridges between other actors or groups of actors. Such central actors control the spread of information between groups of non-adjacent actors. Betweenness centrality, C_B, of an actor, v_i, can be computed as follows [115]:

$$C_B(v_i) = \sum_{j} \sum_{k} \frac{g_{jk}(n_i)}{g_{jk}} \qquad i \neq j \neq k \tag{3}$$

where n is the total number of actors in the social network, $C_B(v_i)$ is the betweenness centrality for actor v_i and g_{jk} is the number of geodesics linking actors v_j and v_k that also pass through actor v_i.

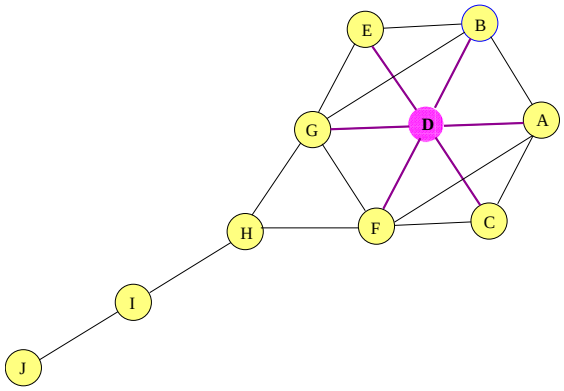
As shown in Figure 4, different central actor(s) in a network can be identified using each of these structural measures (degree, closeness, and betweenness).

In the following, we present how social network analysis and its related structure-based measures have been used in the context of social networks' privacy and list some related privacy threats.

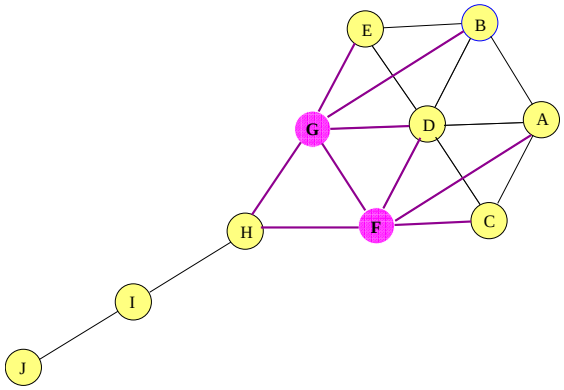
6.2 Privacy Threats

The availability of social network data has attracted the interest of the academic community, third-party advertisers, and governmental services for the purpose of data analysis. Anonymizing these networks before their release is important to enforce privacy but only hiding the identity of the users or removing all their attributes from their profiles does not always guarantee privacy. An attacker can potentially infer the true identities of the targeted users by referring to the structure of the network and by using a background knowledge [93].

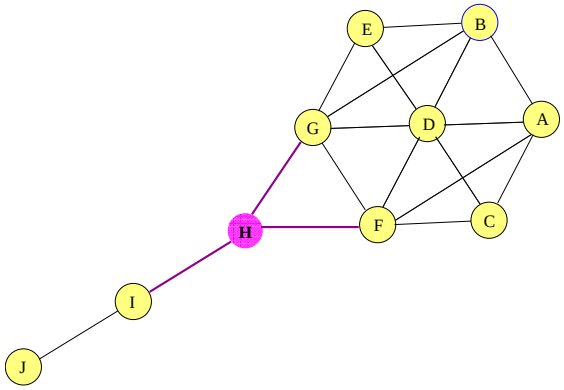
Exploiting structural information with anonymized networks adds a new privacy-related dimension to consider and a large number of theoretical investigations and practical applications have been conducted on social networks [89]



(a) Centrality Degree



(b) Closeness Degree



(c) Betweenness Degree

Fig. 4. A network shaped as a kite graph where each centrality measure yields a different central actor: degree centrality (D), closeness centrality (F and G), and betweenness centrality (H).

[90] [116] [117]. Anonymizing social network data is much more challenging than anonymizing relational data [93]. As stated earlier, social networks can be represented as graphs and thus the social network data can then pre-processed and analyzed via social network analysis measures.

There has been much recent interest to study anonymized networks when an attacker has background information about the network structure [118] [119] [88]. The authors in [118] described a family of attacks based on the structural information of the network. In the first type of attacks, the active attacks, attackers were able to modify the network prior to its release and can potentially construct highly distinguishable subgraphs by inserting nodes and edges to the network. Passive attacks, the other type of attacks, are launched after the anonymized network is published and without inserting new nodes or edges. As observed in [119], the extent to which an individual can be distinguished using graphical position depends on the structural similarity of actors in a network network and the background information an attacker can obtain. More specifically, the structural information is closely related to the degrees of the nodes and their neighbors in a network. In the same line of research, the authors in [88] pointed out that the degree of a node in a graph, among other structural characteristics, can, to a large extent, distinguish the node from other nodes. Consequently, attackers can greatly benefit from structural characteristics of networks that become identifying attributes. An interesting work presented in [120] where the attacks aimed to re-built the network from disparate pieces of information in order to gain more information and better visibility before launching the attacks. The attack consisted in acquiring information about local neighborhoods of different users in the network. Such attack is feasible and its effectiveness depends on the underlying social graph and the degree distribution of its nodes as detailed in the study [120]. The authors concluded that any social network that wishes to enforce the privacy of its users should take great care in decreasing the vulnerability of its interface, i.e. by not displaying the exact number of connections that each users has.

To sum up, social network analysis provides a set of measures that are used extensively to study networks' characteristics and users' behaviors. Though the study of social network is valuable to lot of people, there is serious risk of privacy concerns. Privacy attacks can take advantage of social network analysis to infer further knowledge about social network users using structural information. In addition, the proliferation of online social networks has resulted in huge amounts of available network data. It's quite challenging to protect users' privacy as many sources of background knowledge are widely available on the web. This is particularly true since the majority of social network users are not privacy experts. It is therefore a necessity that social networks enforce users privacy by protecting not only their data but also their established relationships. Moreover, it is highly desirable that social networks implement more flexible and more secure web interfaces. This would greatly benefit users to intuitively configure their privacy settings. At the same time, this would make more difficult for

attackers to use search interfaces, Application Programming Interface (API), or users' connections in order to easily collect user-related information.

7 Link Mining: Tasks and Threats

Owing to the popularity of World Wide Web, the increase of computational power and performance, and the higher capacity to gather and analyze data, large-scale social networks studies are flourishing, spilling over all traditional disciplinary boundaries for social networks. Link mining studies, with their objective to efficiently discover valuable and inherent information from large databases, are highly related to privacy preservation [121]. While centrality measures are widely used in social network analysis [14], link mining techniques rely on recent advances in data mining and often put emphasis on the links between social network actors [15]. In the following, we start by presenting the various link mining tasks before describing relevant privacy threats related to each task.

7.1 Development and Tasks

Taking into account the links between social network actors, various data mining techniques [122] have contributed in the emergence of a new area commonly named *link mining*, where links that exhibit rich patterns are central in extracting hidden knowledge from available data. Link analysis, relational learning, web mining, and graph mining are among the most widely used techniques in link mining [15] [122]. By building predictive models (predicting attributes' values) or descriptive models (extracting interesting patterns), link mining can be regarded as data mining applied on social networks where links play a key role. Not only networks' links can be used to discover prominent actors within a network, but also to reveal uncovered information related to identities, classes, and relationships between actors. In the following, we detail all the tasks that link mining embodies as presented in [15]:

1. **Node-related Approaches:**
   (a) **Link-based Node Ranking [123] [124] [125] [126]:** The objective of link-based node ranking is to prioritize corresponding nodes based on their measured importance. In link mining, centrality measures (e.g., degree, closeness, betweenness, etc.) are used to rank the nodes by exploiting the network structure.
   (b) **Link-based Node Classification [127] [128] [129] [130]:** The link-based node classification task classifies the nodes of a network to a finite set of categories. This type of classification is not only based on nodes' attributes but also on their links to other nodes and on the attributes of these linked nodes.
   (c) **Link-based Node Clustering [131] [132]:** Node clustering, also called group detection, is another well-studied link mining task. Its objective is to identify similar nodes and group them together without predefining

the clusters. Any two nodes, members of the same cluster, are more similar to each other than to any other node in a different cluster. They represent communities where the level of interaction or communication (emails, messages, collaborations, etc.) between nodes of the same cluster is higher than with any node in another cluster.

(d) **Link-based Node Identification [133] [134] [135] [57]:** Link-based node identification, or entity resolution, aims at finding correspondences between nodes of distinct networks given that the nodes that have different identifiers may refer to the same real-world entity. In this case, these nodes form a matched entity pair.

2. **Link-related Approaches:**
   (a) **Link Prediction [136] [137] [138] [139] [140]:** Link prediction, or link existence prediction, is the task of inferring the existence of a link between two nodes, based on the properties of the nodes. While link prediction in static networks aims at inferring missing links and facilitating the task of link formation, link prediction in dynamic networks consists of predicting the snapshot of links at a future time.
   (b) **Link Type Prediction [62] [141]:** Unlike link prediction, where the aim is to predict the existence of a link between two nodes at a particular time, link type prediction aims to identify the type of an existing link (e.g., the type of relationship between two actors). In this task, it is assumed that the existence of a link between the nodes is already confirmed.
3. **Graph-related Approaches:**
   (a) **Subgraph Discovery [142] [143]:** Subgraph discovery is a link mining task that detects similar substructures in pairs of graphs. Its aim is to find the set of subgraphs that are similar among the underlying graphs.
   (b) **Graph Classification [144] [145]:** Graph classification aims at classifying an entire graph with respect to a specific category. Independently classifying each node in a large graph is a tedious task, sometimes infeasible, and may ignore useful information available from other nodes. Rather than trying to label each node within a graph, collective classification learns and infers labels of linked nodes together.
   (c) **Graph-based Generative Models [146] [147]:** Generative models for graphs try to understand the characteristics of networks. Given an input network, generative models can produce a new network similar to the input one. They can model the structure similarities and the data distribution correctly. They are used to model the mechanisms of networks growth and evolution, and to generate networks with realistic properties given few parameters. Studying generative models for graphs is becoming increasingly important, in particular when temporal metrics are considered (e.g., a social network that evolves over time).

As shown in Table 4, link mining techniques have been applied on various networks. Primarily the focus of most link mining approaches has been directed towards the bibliographic [123] [127] [136], biological [129] [135] [138] [144] [147]

[143], and social networks [139] [126] [57] [62] [132] [145] [140] [146]. On these networks, a number of node-related, link-related, and graph-related link mining tasks has been used. Meanwhile, on criminal [131], epidemiology [130], financial [124], and linked data networks [141] [125] [128], node-related techniques have been used. As for link-related approaches, they also examined the data management [133], digital libraries [137], and lexical networks [134]. Besides the biological [143] [144] [147] and social networks [145] [146], graph-related tasks have been applied also on software behavior networks [142]. In the following, we detail a number of privacy threats related to each of the mentioned link mining task.

Table 4. A summary of link mining tasks applied on different types of networks

Network type	Node-related	Link-related	Graph-related
Criminal [131]	✓		
Epidemiology [130]	✓		
Financial [124]	✓		
Linked data [141] [125] [128]	✓		
Database management [133]		✓	
Digital libraries [137]		✓	
Lexical [134]		✓	
Software behavior [142]			✓
Bibliographic [123] [127] [136]	✓	✓	
Biological [129] [135] [138] [144] [147] [143]	✓	✓	✓
Social [139] [126] [57] [62] [132] [145] [140] [146]	✓	✓	✓

7.2 Privacy Threats

Users concerns regarding privacy of personal information are rising in the light of the recent link mining advances. We describe in the following how each link mining task can be exploited by attackers or malicious users. Table 5 lists a number of approaches that may use link mining to compromise users' privacy.

Node-related Threats Social network users have strong expectations of privacy [118] [148]. Tracing users' interactions and reconstructing details of their behaviors are commonly unappreciated. However, *link-based node ranking* can be used to measure the influence and the importance of social network users. Exploiting the structure of the network makes it possible to infer meaningful relationships and quantify the interactions between social network users. Identifying

influential users, who are capable of stimulating other users, is of considerable importance in many scenarios [149]. With the huge number of social network users, link-based node ranking has been applied in many areas such as marketing [150], diffusion of information [151], governmental intelligence-gathering tasks [152].

The privacy implications of *link-based node classification* and *link-based node clustering* can reveal sensitive information such as membership to a particular group or a political party. Node classification and node clustering have been mainly used in the area of computer security. Their extend goes far beyond the simple case of social networks and targeted advertising to reach critical applications such as terrorist networks. Besides the typical profile attributes, users' activities and interactions over time are among the most important sources of information [153]. These online activities and interactions come in many guises such as establishing connections, exchanging messages, and publishing photos.

Link-based node classification can be a source of privacy threat. Prior studies have shown that communities are usually formed around users who share certain interests [154] [155]. Mislove et al. [156] have also reported that social network users are often friends with users who share their attributes. By combining the network graph structure with the fraction of available information, it is possible to infer the value of missing or hidden attributes. For instance, in many situations, some social network users would like to keep private their political affiliations because of privacy concerns. However, node classification techniques can easily infer a user's political affiliation by referring to her contacts and by using, for instance, available information revealing that the user participated in events hosted by a particular political party (e.g., electoral campaigns, debates, etc.) [157].

Link-based node clustering techniques are also a potential source of privacy threat. They are used to group users having the same type of activities, interested in the same hobbies, or seek the same kind of services. Although the user's personal interests are hidden, node clustering techniques can be used to reveal these information using for example some undirect data such as group memberships. In fact, it is possible to extract general groups' interests and then attribute them to users that are members of these groups. In addition, social network groups can be described based on their members' profiles [158]. This example of node clustering is one illustrative example among many that raise users' privacy concerns in social network [159] [160].

Social network users create several accounts on various sites where they disclose personal and professional information [57]. *Link-based node identification* can be used to associate a user profile to a real-world entity (person). Such profile-entity mapping leads to an identity disclosure problem whenever the user would like to keep her social network profile private or hidden. Furthermore, identity disclosure often cause attribute disclosure. Attribute disclosure occurs when sensitive information such as real name, address, and sexual orientation are revealed. In these types of privacy threat, the attacker might have access

to an external knowledge and can use explicit identifiers or quasi-identifiers to reveal the identity of an anonymized user [161] [88].

Link-related Threats *Link prediction* can raise privacy concerns when the predicted link is between users who would like to keep their relationship private. In many cases, the link can be considered as the sensitive information to keep protected. Hiding its existence can be valuable in many real life situations to prevent user-associated sensitive information from being disclosed to third parties [94], to recommend accurately social links without disclosing sensitive information about users' contacts [162], to ensure web browsing anonymity [163]. More interestingly, a recent work shows the possibility to infer whether two non-members friends of a social network member (user) are friends themselves. The obtained results show a high rate of prediction success and is based only on information extracted from the friendship and email contact information of the social network members.

Link type prediction attacks can reveal the sensitive type of an existence relationship between two users although these users would like to keep this information private. Unlike link prediction, link type prediction is concerned in keeping private the type of the relationship (not the existence) between two users. This type of attack is also known as *link re-identification*. This occurs when an attacker is able to identity the type of a sensitive relationship or communication between two users [121]. The type of link between two users can be used to reveal much more information than just the existence of a relationship. Heatherly et al. used link types to classify unknown nodes as terrorist or non-terrorist [164]. In addition, link types can improve the classification accuracy when an attacker attempt to identify information related to personal interests, physical location, political affiliations, etc. For instance, friendship links are more important than professional links to infer personal interests (e.g., political affiliations, religious beliefs, etc.) particularly if a significant number of friends publicly display on their profiles such personal or sensitive information.

Graph-related Threats Attacks based on *subgraph discovery* attempt to acquire new information about an anonymized network using structural subgraph queries. Subgraph queries are useful to efficiently find similar structures in large social networks using two main types of attacks: passive and active attacks [89] [90]. In both types of attacks, structural information is used to reveal the true identities of the targeted users [118]. Consequently subgraph discovery attacks can be used to compromise the privacy of users and their contacts by raising identity and social link disclosure problems [12].

Graph classification which consists at classifying jointly a large number of interconnected nodes in a graph is one interesting aspect of collective classification [15]. In the context of social networks, a number of studies have revealed that users tend to establish friendship ties with other users who have similar interests [154] [155]. This tendency can cover a wide number of sensitive information, such as race and ethnicity, age, religion, education, occupation, etc., which are all

personal attributes. Social network users can set the visibility of their attribute profiles, but this may not be enough to keep private their sensitive information because group memberships and friendship relationships in many cases remain visible and hence can be used to infer private information [165] [159]. In addition, information from the users' contacts can also be extracted from contacts who are not much concerned about securing their personal information or interact. For instance, it is possible to construct automatically users' profiles or infer the values of missing attributes as it is shown in [166] where the authors describe a user profiling approach in social networks. Likewise, other privacy attacks are described in [159] where a mixture of public and private user profiles are used to predict the private attributes of users by applying collective classification which aims at learning and inferring class labels of linked nodes together.

Understanding the structure and evolution of social networks over time have gained much attention recently, especially with many *graph-based generative models* [167] [168] [169]. The potential of these approaches to study network evolution and group formation is appealing and, as a consequence, the insights provided by these methods are highly interesting: How networks and group are formed? Why and when users join groups? Which groups will grow rapidly? At the same time, the consequences are rather severe for users' privacy when these approaches are misused. For instance, the authors in [168] address the problem of modeling social network generation and demonstrate the capacity of their generative model to reveal that users are joining groups for various reasons and that friendship with other group members is only one of these reasons. In another work [64], the author propose a generative model for friendships in social circles using a combination of both network and profile information. This model can automatically identify users' social circles (e.g., contacts who are friends, contacts from the same hometown, contacts from the same college) and predict to which circles a new contact should be assigned.

Table 5. Link mining tasks and their corresponding implications

Technique	**Application and Threat**
Link-based node ranking	[149] [150] [151] [152]
Link-based node classification	[156] [157]
Link-based node clustering	[158] [159] [160]
Link-based node identification	[57] [161] [88]
Link prediction	[94] [162] [163]
Link type prediction	[121] [164]
Subgraph discovery	[89] [90] [118] [12]
Graph classification	[165] [159] [166]
Generative models	[168] [64]

To conclude, link mining techniques can be used to infer new information about the users and consequently pose serious privacy concerns. This can take

many forms within the different link mining tasks. In node-related tasks attackers can compromise users' privacy by exploring the structure of social networks, identifying key users, combining information about the network structure with the fraction of publicly available information, investigating group memberships, and associating users' profiles to their corresponding real-world entities. In link-related tasks, attackers can invade users' privacy by predicting link existence between users who would like to keep their relationship private, and revealing the type of a sensitive relationship between users. In graph-related tasks, attackers can violate users' privacy by re-identifying users of two similar structures (an anonymized network and a background of knowledge), classifying jointly a large number of interconnected nodes, extracting information from contacts with public profiles, and seeking to retrieve insights from the study of network evolution and group formation. Concerned about privacy implications, social network users are increasingly interested in solutions that enforce individual privacy and protect sensitive information. Next, we discuss some open challenges that future social network solutions must take into account.

8 Open Challenges

As information sharing is gaining a great deal of attention among social networks users, privacy on these sites is one of the most intriguing social networks' challenges [9]. Data on social networks usually open up questions related to users' privacy and data management. Privacy concerns vary significantly across these networks due to the open nature of how data is displayed and controlled by each site. In the following, we present some of the key challenges to consider in future privacy protection approaches.

8.1 Relationship Discovery and Management

Currently, social networks let users manage and control their privacy settings. However, they typically do so in terms of contact identities or manually grouped lists of contacts. Controlling access to own resources is driven more by relationships that social network users share with their contacts such as colleagues, relatives, friends, etc. Treating all their contacts in the same way, without differentiating one user from another, is an unsafe and restrictive practice. For instance, a user might want to:

1. Prevent her relatives from viewing content posted by her friends,
2. Prevent her relatives from connecting, posting comments, and communicating with her colleagues,
3. Share some data and interact with some members of their social network but not with all of them.

Unfortunately, relationship types between a user and her contacts are often missing [36]. Although some social network sites provide the possibility to define

manually how a user knows each of her contacts, most of the time this option is skipped by social network users and only the link existence is indicated [170].

Relationship discovery and management is one of the key challenges that are closely related to privacy topics. On social networks where relationships multiply rapidly and evolve over time, relationship-based management mechanisms play a major role in enforcing and facilitating privacy-related settings. This involves mechanisms that let social network users protect their personal information by only granting access to some contacts (e.g., friends) while denying the access to other contacts (e.g., colleagues). Social network users are facing different kinds of misuse cases regarding their privacy due to the lack of efficient access control models. While relationship-based access control mechanisms seems highly interesting, such approaches are rarely implemented since relationship types on social networks remain unlabeled. Recently, the problem of missing relationship types in social networks has been investigated in [62] [64] where a number of relationship discovery approaches are described. Such relationship-related capacities seem to be promising when approaching privacy issues, and social networks should explicitly consider them in their privacy settings far more than they currently do.

8.2 Multimedia Content and Metadata

Sharing and controlling access to published information has become an integral part of users' every day lives. Apparently, users are cautious about the information they reveal online in terms of [171]:

1. Degree of identifiability (pseudonyms/real names),
2. Type of information (school name, hobbies, interests, etc.),
3. Visibility of information (who can view user's profile).

At the same time, the number of created accounts on various sites and the volume of data available on social networks are exponentially growing at an incredible rate, going beyond social network users' ability to easily interact and manage their data. In addition, users are unaware of the risks associated with the disclosed information since they cannot fully control who can use these information and for what purposes [172].

Managing multimedia content and their metadata is another challenge for social network users. This particular privacy challenge is raised by the increasingly growing number of multimedia objects uploaded on social networks. For instance, 250 million photos are uploaded on Facebook[19] every day. However, photos can be used to identify persons (e.g., facial recognition), as well as to infer additional information through available metadata (e.g., GPS location, data and time, device name, etc.). Personal photos published on social networks may be used for inappropriate purposes related to users' private life, friends, work information, habits, etc. For example, employers who want to justify a decision to fire

[19] https://www.facebook.com/press/info.php?statistics, accessed 01 October 2012

an employee or to check the backgrounds of potential employee. Similarly, users' friends or the friends of their friends who are interested to know more about the past of a person or infer new information more than what the users chose to share with them (list of friends, physical location, political affiliation, sexual orientation, etc.). More importantly, such available information when associated with other profile content, may be used by family members, friends, or colleagues to check information related to sexuality, relationship status, location, or details of personal problems that owners might consider embarrassing if it was widely known [173].

It is therefore obvious that unauthorized users or parties must not be able to link between the various user's activities in order to infer further information. Hiding the link between multiple actions is essential to prevent attackers to reconstruct user's profile. This aspect is essentially related to the unlinkability privacy requirement. Novel approaches must ensure that private data must be protected and no useful information can be leaked through the analysis of user's activities.

8.3 Social Media Preservation

There is no doubt that social networks sites can yield much more information from users' data such as selling insights from mined data for targeted advertising purposes [174], predicting relationships from social behavior data useful in understanding topics discussed between users and their personalities [153], and falsifying identities for criminal and terrorist groups [175]. At any time and often without users' knowledge, social networks can mine, copy, or archive personal information. Personal data can be mined and used to reveal information about users' private lives, their social relationships, or additional information that users would like to keep private. Unaware of an old posted information or a previously added online contact, an uncountable number of social network users went through bad experiences that affected their life.

Dynamically protecting sensitive information of the archived users' data is another privacy challenge on social networks. Social network sites have complete control over users' data and may intentionally or accidentally leak its content to unauthorized entities or third parties. A fundamental privacy concern for users is that social network providers and third-parties may potentially access and aggregate personal information from the users' archived data [176] [10] [177]. However, the social network users' perception of privacy goes beyond hiding information from being viewed by their contacts. It also involves mechanisms that enforce privacy and ensure that personal archived information is not misused. And if information is leaked, even though appropriate mechanisms were implemented, an enforced privacy must be able to guarantee that no personal information is revealed to unauthorized entities. This privacy aspect underlies the unobservability aspect of privacy. While many social networks provide various forms of access control mechanisms to restrict who can view a published information, they do not provide any form of protection regarding the possibility that archived information can be misused and consequently analyzed to reveal

sensitive information. To date, none of these concerns has been successfully integrated into users' privacy settings, and the privacy-related issues derived from social media preservation have been largely left unexplored.

8.4 Social Digital Ecosystem

Along with the previously mentioned privacy challenges, future social networks must shed the lights on the real needs and expectations of their users. So far, social network providers set the boundaries within which users can socialize, share information, accept data ownership rights, permission policies, as well as other critical issues. Current social networks dictate such rules rather than letting their users set up their own rules for data sharing and ownerships to promote data decentralization and to go beyond the walled garden of the current social network sites.

Shifting social networks toward an ecosystem model is the challenge for the new generation of social networks. In essence, an ecosystem is an environment made of entities that interact within the system, maintain the system stable, are committed to ensure mutual respect, and can benefit from each other's participation [178]. Shifting the future forms of social networks toward an ecosystem environment would enable users to set their own preferences and to use the system more effectively. These preferences are initially the same for all social network users. After joining an ecosystem, users are free to personalize these preferences as they wish. One aspect of these settings is related to the privacy of the users who can design their privacy strategies as they wish. This would ensure a better protection level since all the social ecosystem are enforced to accept the *gold rule* that englobes the system's stability, the mutual respect, and the positive participation of the users.

9 Conclusion

This chapter has presented a global overview about privacy on social networks. While it is complex to give a precise definition of privacy as it is engorged with various and distinct meanings, understanding the characteristics of social networks would certainly contribute positively to the design of more adequate privacy protection approaches.

As stated earlier, by analyzing social networks' structure and content it is possible to infer further knowledge which may go beyond what the users want to disclose. In light of this, we reviewed various social network analysis measures and link mining techniques then we derived their associated privacy threats. Particularly, we have shown the importance of taking into consideration these techniques and the usefulness to integrate them in future privacy-related systems. Novel approaches must be able to cope with the various privacy challenges such as trust and privacy management, risks and threats of social networking, traceability analysis, user profiling and related risks, ethical conflicts in social

networks as well as the moral implications, relationship management and discovery, anonymity preserving, social terrorism, social network-based access control, and abnormal activities on social networks.

Of all social network challenges, privacy protections is crucial for both users and social networks. Failing to provide an optimal privacy protection may have undesirable consequences on the popularity of such social networks and on the amount of information that social network users are willing to share. It is therefore up to social network sites to provide their users with a variety of support tools that align with users' perceptions of privacy such as enhanced relationship management capacities, intuitive interfaces, fine-grained access control, secure online data storage, and media preservation solutions.

References

1. Liu, Y., Gummadi, K.P., Krishnamurthy, B., Mislove, A.: Analyzing facebook privacy settings: user expectations vs. reality. In: Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. IMC '11, ACM (2011) 61–70
2. Lipford, H.R., Besmer, A., Watson, J.: Understanding privacy settings in facebook with an audience view. In: Proceedings of the 1st Conference on Usability, Psychology, and Security, USENIX Association Berkeley, CA, USA (2008) 1–8
3. Acquisti, A., Gross, R.: Imagined communities: Awareness, information sharing, and privacy on the facebook. In Danezis, G., Golle, P., eds.: Privacy Enhancing Technologies. Volume 4258 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2006) 36–58
4. Awad, N., Krishnan, M.: The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. MIS quarterly (2006) 13–28
5. Madejski, M., Johnson, M., Bellovin, S.: A study of privacy settings errors in an online social network. In: Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on. (2012) 340–345
6. Johnson, M., Egelman, S., Bellovin, S.M.: Facebook and privacy: it's complicated. In: Proceedings of the Eighth Symposium on Usable Privacy and Security. SOUPS '12, ACM (2012) 1–15
7. Besmer, A., Richter Lipford, H.: Moving beyond untagging: photo privacy in a tagged world. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '10, ACM (2010) 1563–1572
8. Grimmelmann, J.: Saving facebook. Iowa Law Review **94** (2009) 1137–1206
9. Zhang, C., Sun, J., Zhu, X., Fang, Y.: Privacy and security for online social networks: challenges and opportunities. IEEE Network **24**(4) (july-august 2010) 13–18
10. Viswanath, B., Kiciman, E., Saroiu, S.: Keeping information safe from social networking apps. In: Proceedings of the 2012 ACM workshop on Workshop on online social networks. WOSN '12, ACM (2012) 49–54
11. Brey, P.: Ethical aspects of information security and privacy. In Petkovic, M., Jonker, W., Carey, M.J., Ceri, S., eds.: Security, Privacy, and Trust in Modern Data Management. Data-Centric Systems and Applications. Springer Berlin Heidelberg (2007) 21–36

12. Zheleva, E., Terzi, E., Getoor, L.: Privacy in social networks. Synthesis Lectures on Data Mining and Knowledge Discovery **3**(1) (2012) 1–85
13. Pfitzmann, A., Hansen, M.: A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management (August 2010) v0.34.
14. Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press (1994)
15. Getoor, L., Diehl, C.P.: Link mining: a survey. SIGKDD Explor. Newsl. **7**(2) (December 2005) 3–12
16. Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., Wiener, J.: Graph structure in the web. Computer Networks **33**(1) (2000) 309–320
17. Wellman, B.: Computer networks as social networks. Science **293**(5537) (2001) 2031–2034
18. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., Sakaki, Y.: A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proceedings of the National Academy of Sciences **98**(8) (2001) 4569–4574
19. Gomez-Rodriguez, M., Leskovec, J., Krause, A.: Inferring networks of diffusion and influence. ACM Transactions on Knowledge Discovery from Data **5**(4) (2012)
20. Adamic, L., Buyukkokten, O., Adar, E.: A social network caught in the web. First Monday **8**(6) (2003)
21. Wilson, C., Sala, A., Puttaswamy, K.P.N., Zhao, B.Y.: Beyond social graphs: User interactions in online social networks and their implications. ACM Trans. Web **6**(4) (November 2012) 17:1–17:31
22. Faust, K.: Centrality in affiliation networks. Social Networks **19**(2) (1997) 157 – 191
23. Borgatti, S., Mehra, A., Brass, D., Labianca, G.: Network analysis in the social sciences. Science **323**(5916) (2009) 892–895
24. Matsuo, Y., Mori, J., Hamasaki, M., Nishimura, T., Takeda, H., Hasida, K., Ishizuka, M.: Polyphonet: An advanced social network extraction system from the web. Web Semantics **5**(4) (2007) 262–278
25. Gilbert, E.: Predicting tie strength in a new medium. In: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. CSCW '12, ACM (2012) 1047–1056
26. Haythornthwaite, C.: Social network analysis: An approach and technique for the study of information exchange. Library and Information Science Research **18**(4) (1996) 323–342
27. Musial, K., Kazienko, P.: Social networks on the internet. World Wide Web **16**(1) (2013) 31–72
28. Granovetter, M.S.: The strength of weak ties. American Journal of Sociology **78**(6) (1973) 1360–1380
29. Haythornthwaite, C.: Social networks and internet connectivity effects. Information, Communication & Society **8**(2) (2005) 125–147
30. Wellman, B., Wortley, S.: Different strokes from different folks: Community ties and social support. American Journal of Sociology **96**(3) (1990) 558–588
31. Bernard, H., Johnsen, E., Killworth, P., McCarty, C., Shelley, G., Robinson, S.: Comparing four different methods for measuring personal social networks. Social Networks **12**(3) (1990) 179–215
32. Killworth, P., Johnsen, E., Bernard, H., Ann Shelley, G., McCarty, C.: Estimating the size of personal networks. Social Networks **12**(4) (1990) 289–312

33. Steinfield, C., Ellison, N., Lampe, C.: Social capital, self-esteem, and use of online social network sites: A longitudinal analysis. Journal of Applied Developmental Psychology **29**(6) (2008) 434–445
34. Heidemann, J., Klier, M., Probst, F.: Online social networks: A survey of a global phenomenon. Computer Networks **56**(18) (2012) 3866–3878
35. Hua, W., Wellman, B.: Social connectivity in america: Changes in adult friendship network size from 2002 to 2007. American Behavioral Scientist **53**(8) (2010) 1148–1169
36. Dunbar, R.I.M.: Social cognition on the internet: testing constraints on social network size. Philosophical Transactions of the Royal Society B: Biological Sciences **367**(1599) (2012) 2192–2201
37. Subrahmanyam, K., Reich, S., Waechter, N., Espinoza, G.: Online and offline social networks: Use of social networking sites by emerging adults. Journal of Applied Developmental Psychology **29**(6) (2008) 420–433
38. Correa, T., Hinsley, A., de Zúñiga, H.: Who interacts on the web?: The intersection of users' personality and social media use. Computers in Human Behavior **26**(2) (2010) 247–253
39. Joinson, A.N.: Looking at, looking up or keeping up with people?: motives and use of facebook. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '08, ACM (2008) 1027–1036
40. Papacharissi, Z., Mendelson, A.: Toward a new (er) sociability: uses, gratifications and social capital on facebook. Media perspectives for the 21st century (2011) 212–230
41. Ryan, T., Xenos, S.: Who uses facebook? an investigation into the relationship between the big five, shyness, narcissism, loneliness, and facebook usage. Computers in Human Behavior **27**(5) (2011) 1658–1664
42. Gentile, B., Twenge, J., Freeman, E., Campbell, W.: The effect of social networking websites on positive self-views: An experimental investigation. Computers in Human Behavior **28**(5) (2012) 1929–1933
43. Nosko, A., Wood, E., Molema, S.: All about me: Disclosure in online social networking profiles: The case of facebook. Computers in Human Behavior **26**(3) (2010) 406–418
44. Krasnova, H., Spiekermann, S., Koroleva, K., Hildebrand, T.: Online social networks: Why we disclose. Journal of Information Technology **25**(2) (2010) 109–125
45. Ross, C., Orr, E., Sisic, M., Arseneault, J., Simmering, M., Orr, R.: Personality and motivations associated with facebook use. Computers in Human Behavior **25**(2) (2009) 578–586
46. Morris, M.R., Teevan, J., Panovich, K.: What do people ask their social networks, and why?: a survey study of status message q&a behavior. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '10, ACM (2010) 1739–1748
47. Lin, K.Y., Lu, H.P.: Why people use social networking sites: An empirical study integrating network externalities and motivation theory. Computers in Human Behavior **27**(3) (2011) 1152–1161
48. Nosko, A., Wood, E., Kenney, M., Archer, K., De Pasquale, D., Molema, S., Zivcakova, L.: Examining priming and gender as a means to reduce risk in a social networking context: Can stories change disclosure and privacy setting use when personal profiles are constructed? Computers in Human Behavior **28**(6) (2012) 2067–2074

49. Sledgianowski, D., Kulviwat, S.: Using social network sites: The effects of playfulness, critical mass and trust in a hedonic context. Journal of Computer Information Systems **49**(4) (2009) 74–83
50. Pempek, T., Yermolayeva, Y., Calvert, S.: College students' social networking experiences on facebook. Journal of Applied Developmental Psychology **30**(3) (2009) 227–238
51. Kwon, O., Wen, Y.: An empirical study of the factors affecting social network service use. Computers in Human Behavior **26**(2) (2010) 254–263
52. Adamic, L., Adar, E.: How to search a social network. Social Networks **27**(3) (2005) 187–203
53. Boyd, D., Ellison, N.: Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication **13**(1) (2007) 210–230
54. Schneider, F., Feldmann, A., Krishnamurthy, B., Willinger, W.: Understanding online social network usage from a network perspective. In: Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference. IMC '09, ACM (2009) 35–48
55. Bakshy, E., Rosenn, I., Marlow, C., Adamic, L.: The role of social networks in information diffusion. WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web (2012) 519–528
56. Ellison, N., Steinfield, C., Lampe, C.: The benefits of facebook "friends:" social capital and college students' use of online social network sites. Journal of Computer-Mediated Communication **12**(4) (2007) 1143–1168
57. Raad, E., Chbeir, R., Dipanda, A.: User profile matching in social networks. Proceedings - 13th International Conference on Network-Based Information Systems, NBiS 2010 (sept. 2010) 297–304
58. Thelwall, M.: Social networks, gender, and friending: An analysis of mySpace member profiles. Journal of the American Society for Information Science and Technology **59**(8) (2008) 1321–1330
59. Abel, F., Henze, N., Herder, E., Krause, D.: Interweaving public user profiles on the web. In Bra, P., Kobsa, A., Chin, D., eds.: User Modeling, Adaptation, and Personalization. Volume 6075 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2010) 16–27
60. Graves, M., Constabaris, A., Brickley, D.: FOAF: Connecting people on the Semantic Web. Cataloging and Classification Quarterly **43**(3-4) (2007) 191–202
61. Cranshaw, J., Toch, E., Hong, J., Kittur, A., Sadeh, N.: Bridging the gap between physical location and online social networks. UbiComp'10 - Proceedings of the 2010 ACM Conference on Ubiquitous Computing (2010) 119–128
62. Raad, E., Chbeir, R., Dipanda, A.: Discovering relationship types between users using profiles and shared photos in a social network. Multimedia Tools and Applications **64**(1) (2013) 141–170
63. Tang, L., Liu, H.: Scalable learning of collective behavior based on sparse social dimensions. In: Proceedings of the 18th ACM conference on Information and knowledge management. CIKM '09, ACM (2009) 1107–1116
64. McAuley, J., Leskovec, J.: Learning to discover social circles in ego networks. Advances in Neural Information Processing Systems 25 (2012) 548–556
65. Ellison, N., Steinfield, C., Lampe, C.: Connection strategies: Social capital implications of facebook-enabled communication practices. New Media and Society **13**(6) (2011) 873–892
66. Newman, M.: The structure and function of complex networks. SIAM Review **45**(2) (2003) 167–256

67. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U.: Complex networks: Structure and dynamics. Physics Reports **424**(4-5) (2006) 175–308
68. Fortunato, S.: Community detection in graphs. Physics Reports **486**(3-5) (2010) 75–174
69. Brams, S., Mutlu, H., Ramirez, S.: Influence in terrorist networks: From undirected to directed graphs. Studies in Conflict and Terrorism **29**(7) (2006) 703–718
70. Morselli, C., Giguère, C., Petit, K.: The efficiency/security trade-off in criminal networks. Social Networks **29**(1) (2007) 143–153
71. Bachi, G., Coscia, M., Monreale, A., Giannotti, F.: Classifying trust/distrust relationships in online social networks. In: Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom). (sept. 2012) 552 –557
72. Zhang, T., Chao, H., Tretter, D.: Dynamic estimation of family relations from photos. In Lee, K.T., Tsai, W.H., Liao, H.Y., Chen, T., Hsieh, J.W., Tseng, C.C., eds.: Advances in Multimedia Modeling. Volume 6524 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2011) 65–76
73. Carminati, B., Ferrari, E., Perego, A.: Enforcing access control in web-based social networks. ACM Transactions on Information and System Security **13**(1) (2009)
74. Xiang, R., Neville, J., Rogati, M.: Modeling relationship strength in online social networks. In: Proceedings of the 19th international conference on World Wide Web. WWW '10, ACM (2010) 981–990
75. Stutzman, F., Kramer-Duffield, J.: Friends only: examining a privacy-enhancing behavior in facebook. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '10, ACM (2010) 1553–1562
76. KogovŽek, T., Ferligoj, A., Coenders, G., Saris, W.: Estimating the reliability and validity of personal support measures: Full information ML estimation with planned incomplete data. Social Networks **24**(1) (2002) 1–20
77. Groves, R.: Survey errors and survey costs. Volume 536. Wiley-Interscience (2004)
78. Gonzalez-Bailon, S.: Opening the black box of link formation: Social factors underlying the structure of the web. Social Networks **31**(4) (2009) 271–280
79. Tyler, J., Wilkinson, D., Huberman, B.: E-mail as spectroscopy: Automated discovery of community structure within organizations. Information Society **21**(2) (2005) 133–141
80. Kautz, H., Selman, B., Shah, M.: Referral web: combining social networks and collaborative filtering. Commun. ACM **40**(3) (March 1997) 63–65
81. Mika, P.: Flink: Semantic web technology for the extraction and analysis of social networks. Web Semantics **3**(2-3) (2005) 211–223
82. Marin, A., Wellman, B.: Social network analysis: An introduction. Handbook of Social Network Analysis **22**(January) (2010) 11–25
83. Kwak, H., Lee, C., Park, H., Moon, S.: What is twitter, a social network or a news media? In: Proceedings of the 19th international conference on World Wide Web. WWW '10, ACM (2010) 591–600
84. Kazienko, P., Musial, K., Kajdanowicz, T.: Multidimensional social network in the social recommender system. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on **41**(4) (2011) 746–759
85. Catanese, S., Meo, P., Ferrara, E., Fiumara, G., Provetti, A.: Extraction and analysis of facebook friendship relations. In Abraham, A., ed.: Computational Social Networks. Springer London (2012) 291–324
86. Schneier, B.: A taxonomy of social networking data. IEEE Security and Privacy **8**(4) (2010) 88

87. Diesner, J., Frantz, T., Carley, K.: Communication networks from the enron email corpus "it's always about the people. enron is no different". Computational and Mathematical Organization Theory **11**(3) (2005) 201–228
88. Liu, K., Terzi, E.: Towards identity anonymization on graphs. In: Proceedings of the 2008 ACM SIGMOD international conference on Management of data. SIGMOD '08, ACM (2008) 93–106
89. Narayanan, A., Shmatikov, V.: De-anonymizing social networks. In: Security and Privacy, 2009 30th IEEE Symposium on. (may 2009) 173 –187
90. Hay, M., Miklau, G., Jensen, D., Towsley, D., Li, C.: Resisting structural re-identification in anonymized social networks. VLDB Journal **19**(6) (2010) 797–823
91. Agrawal, R., Srikant, R.: Privacy-preserving data mining. SIGMOD Record (ACM Special Interest Group on Management of Data) **29**(2) (2000) 439–450
92. Verykios, V., Bertino, E., Fovino, I., Provenza, L., Saygin, Y., Theodoridis, Y.: State-of-the-art in privacy preserving data mining. SIGMOD Record **33**(1) (2004) 50–57
93. Zhou, B., Pei, J.: The k-anonymity and l-diversity approaches for privacy preservation in social networks against neighborhood attacks. Knowledge and Information Systems **28**(1) (2011) 47–77
94. Li, N., Zhang, N., Das, S.K.: Relationship privacy preservation in publishing online social networks. In: PASSAT/SocialCom. (oct. 2011) 443 –450
95. Fong, P.W., Siahaan, I.: Relationship-based access control policies and their policy languages. In: Proceedings of the 16th ACM symposium on Access control models and technologies. SACMAT '11, ACM (2011) 51–60
96. Cheek, G.P., Shehab, M.: Policy-by-example for online social networks. In: Proceedings of the 17th ACM symposium on Access Control Models and Technologies. SACMAT '12, ACM (2012) 23–32
97. Li, N., Chen, G.: Sharing location in online social networks. IEEE Network **24**(5) (2010) 20–25
98. Backstrom, L., Sun, E., Marlow, C.: Find me if you can: improving geographical prediction with social and spatial proximity. In: Proceedings of the 19th international conference on World Wide Web. WWW '10, ACM (2010) 61–70
99. Cunningham, S., Masoodian, M., Adams, A.: Privacy issues for online personal photograph collections. Journal of Theoretical and Applied Electronic Commerce Research **5**(2) (2010) 26–40
100. Frikken, K., Atallah, M., Li, J.: Attribute-based access control with hidden policies and hidden credentials. IEEE Transactions on Computers **55**(10) (2006) 1259–1270
101. Bouna, B., Chbeir, R., Marrara, S.: Enforcing role based access control model with multimedia signatures. Journal of Systems Architecture **55**(4) (2009) 264–274
102. Peng, H., Gu, J., Ye, X.: Dynamic purpose-based access control. In: Parallel and Distributed Processing with Applications, 2008. ISPA '08. International Symposium on. (2008) 695–700
103. Willinger, W., Rejaie, R., Torkjazi, M., Valafar, M., Maggioni, M.: Research on online social networks: time to face the real challenges. SIGMETRICS Perform. Eval. Rev. **37**(3) (January 2010) 49–54
104. Wellman, B.: The school child's choice of companions. The Journal of Educational Research **14**(2) (1926) 126–132
105. Bott, H.: Observation of play activities in a nursery school. Genetic Psychology Monographs **4**(1) (1928) 44–88

106. Moreno, J.: Who shall survive? Volume 58. Nervous and Mental Disease Publishing Company Washington, DC (1934)
107. Wellman, B.: Network analysis: Some basic principles. Sociological theory **1**(1) (1983) 155–200
108. Barnes, J.A.: Class and committees in a norwegian island parish. Human Relations **7**(1) (1954) 39–58
109. Luke, D., Harris, J.: Network analysis in public health: History, methods, and applications. Annual Review of Public Health **28** (2007) 69–93
110. Watts, D.: The "new" science of networks. Annual Review of Sociology **30** (2004) 243–270
111. Barabási, A.L., Bonabeau, E.: Scale-free networks. Scientific American **288**(5) (2003) 60–69
112. Christley, R., Pinchbeck, G., Bowers, R., Clancy, D., French, N., Bennett, R., Turner, J.: Infection in social networks: Using network analysis to identify high-risk individuals. American Journal of Epidemiology **162**(10) (2005) 1024–1031
113. Mislove, A., Marcon, M., Gummadi, K.P., Druschel, P., Bhattacharjee, B.: Measurement and analysis of online social networks. In: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement. IMC '07, ACM (2007) 29–42
114. Koschützki, D., Lehmann, K., Peeters, L., Richter, S., Tenfelde-Podehl, D., Zlotowski, O.: Centrality indices. In Brandes, U., Erlebach, T., eds.: Network Analysis. Volume 3418 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2005) 16–61
115. Freeman, L.: Centrality in social networks conceptual clarification. Social Networks **1**(3) (1978) 215–239
116. He, X., Vaidya, J., Shafiq, B., Adam, N., Atluri, V.: Preserving privacy in social networks: A structure-aware approach. In: Web Intelligence and Intelligent Agent Technologies. Volume 1. (sept. 2009) 647 –654
117. Campan, A., Truta, T.: Data and structural k-anonymity in social networks. In Bonchi, F., Ferrari, E., Jiang, W., Malin, B., eds.: Privacy, Security, and Trust in KDD. Volume 5456 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2009) 33–54
118. Backstrom, L., Dwork, C., Kleinberg, J.: Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography. In: Proceedings of the 16th international conference on World Wide Web. WWW '07, ACM (2007) 181–190
119. Hay, M., Miklau, G., Jensen, D., Weis, P., Srivastava, S.: Anonymizing social networks. Technical report (2007)
120. Korolova, A., Motwani, R., Nabar, S.U., Xu, Y.: Link privacy in social networks. In: Proceedings of the 17th ACM conference on Information and knowledge management. CIKM '08, ACM (2008) 289–298
121. Zheleva, E., Getoor, L.: Preserving the privacy of sensitive relationships in graph data. In Bonchi, F., Ferrari, E., Malin, B., Saygin, Y., eds.: Privacy, Security, and Trust in KDD. Volume 4890 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2008) 153–171
122. Wu, X., Kumar, V., Ross, Q., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G., Ng, A., Liu, B., Yu, P., Zhou, Z.H., Steinbach, M., Hand, D., Steinberg, D.: Top 10 algorithms in data mining. Knowledge and Information Systems **14**(1) (2008) 1–37
123. Liu, X., Bollen, J., Nelson, M., Van De Sompel, H.: Co-authorship networks in the digital library research community. Information Processing and Management **41**(6) (2005) 1462–1480

124. Creamer, G., Stolfo, S.: A link mining algorithm for earnings forecast and trading. Data Mining and Knowledge Discovery **18**(3) (2009) 419–445
125. Li, P., Li, Z., Liu, H., He, J., Du, X.: Using link-based content analysis to measure document similarity effectively. In Li, Q., Feng, L., Pei, J., Wang, S., Zhou, X., Zhu, Q.M., eds.: Advances in Data and Web Management. Volume 5446 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2009) 455–467
126. Lin, Z., Wang, L., Guo, S.: Recommendations on social network sites: From link mining perspective. In: Management and Service Science, 2009. MASS '09. International Conference on. (sept. 2009) 1 –4
127. Karamon, J., Matsuo, Y., Yamamoto, H., Ishizuka, M.: Generating social network features for link-based classification. In Kok, J., Koronacki, J., Lopez de Mantaras, R., Matwin, S., Mladenic, D., Skowron, A., eds.: Knowledge Discovery in Databases: PKDD 2007. Volume 4702 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2007) 127–139
128. Chakrabarti, S., Dom, B., Indyk, P.: Enhanced hypertext categorization using hyperlinks. SIGMOD Record **27**(2) (1998) 307–318
129. Segal, E., Wang, H., Koller, D.: Discovering molecular pathways from protein interaction and gene expression data. Bioinformatics **19**(SUPPL. 1) (2003) i264–i272
130. Stattner, E., Vidot, N.: Social network analysis in epidemiology: Current trends and perspectives. In: Research Challenges in Information Science (RCIS), 2011 Fifth International Conference on. (may 2011) 1 –11
131. Fard, A., Ester, M.: Collaborative mining in multiple social networks data for criminal group discovery. In: Computational Science and Engineering, 2009. CSE '09. International Conference on. Volume 4. (aug. 2009) 582 –587
132. Barbier, G., Liu, H.: Data mining in social media. In Aggarwal, C.C., ed.: Social Network Data Analytics. Springer US (2011) 327–352
133. Bhattacharya, I., Getoor, L.: Iterative record linkage for cleaning and integration. In: Proceedings of the 9th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery. DMKD '04, ACM (2004) 11–18
134. Ponzetto, S.P., Strube, M.: Exploiting semantic role labeling, wordnet and wikipedia for coreference resolution. In: Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics. HLT-NAACL '06, Association for Computational Linguistics (2006) 192–199
135. Stein, L.: Integrating biological databases. Nature Reviews Genetics **4**(5) (2003) 337–345
136. Taskar, B., Wong, M.F., Abbeel, P., Koller, D.: Link prediction in relational data. In: Advances in Neural Information Processing Systems (NIPS), Cambridge, MA: MIT Press (2003)
137. Huang, Z., Li, X., Chen, H.: Link prediction approach to collaborative filtering. In: Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries. JCDL '05, ACM (2005) 141–142
138. Yu, H., Paccanaro, A., Trifonov, V., Gerstein, M.: Predicting interactions in protein networks by completing defective cliques. Bioinformatics **22**(7) (2006) 823–829
139. Zheleva, E., Getoor, L., Golbeck, J., Kuter, U.: Using friendship ties and family circles for link prediction. Lecture Notes in Computer Science **5498 LNAI** (2009) 97–113

140. Buccafurri, F., Lax, G., Nocera, A., Ursino, D.: Discovering links among social networks. In Flach, P., Bie, T., Cristianini, N., eds.: Machine Learning and Knowledge Discovery in Databases. Volume 7524 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2012) 467–482
141. Getoor, L.: Link mining: a new data mining challenge. SIGKDD Explor. Newsl. **5**(1) (July 2003) 84–89
142. Cheng, H., Lo, D., Zhou, Y., Wang, X., Yan, X.: Identifying bug signatures using discriminative graph mining. In: Proceedings of the eighteenth international symposium on Software testing and analysis. ISSTA '09, ACM (2009) 141–152
143. Ciriello, G., Guerra, C.: A review on models and algorithms for motif discovery in protein-protein interaction networks. Briefings in Functional Genomics (2008)
144. Borgwardt, K., Ong, C., Schönauer, S., Vishwanathan, S., Smola, A., Kriegel, H.P.: Protein function prediction via graph kernels. Bioinformatics **21**(SUPPL. 1) (2005) i47–i56
145. Rabelo, J., Prudêncio, R., Barros, F.: Leveraging relationships in social networks for sentiment analysis. In: Proceedings of the 18th Brazilian symposium on Multimedia and the web. WebMedia '12, ACM (2012) 181–188
146. Zhou, D., Manavoglu, E., Li, J., Giles, C.L., Zha, H.: Probabilistic models for discovering e-communities. In: Proceedings of the 15th international conference on World Wide Web. WWW '06, ACM (2006) 173–182
147. Nguyen, C., Mamitsuka, H.: Kernels for link prediction with latent feature models. In Gunopulos, D., Hofmann, T., Malerba, D., Vazirgiannis, M., eds.: Machine Learning and Knowledge Discovery in Databases. Volume 6912 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2011) 517–532
148. Zhou, B., Pei, J., Luk, W.: A brief survey on anonymization techniques for privacy preserving publishing of social network data. SIGKDD Explor. Newsl. **10**(2) (December 2008) 12–22
149. Bakshy, E., Hofman, J.M., Mason, W.A., Watts, D.J.: Everyone's an influencer: quantifying influence on twitter. In: Proceedings of the fourth ACM international conference on Web search and data mining. WSDM '11, ACM (2011) 65–74
150. Chen, W., Wang, C., Wang, Y.: Scalable influence maximization for prevalent viral marketing in large-scale social networks. In: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. KDD '10, ACM (2010) 1029–1038
151. Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. KDD '03, ACM (2003) 137–146
152. National Research Council (US). Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals: Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment. National Academies Press (2008)
153. Adali, S., Sisenda, F., Magdon-Ismail, M.: Actions speak as loud as words: predicting relationships from social behavior data. In: Proceedings of the 21st international conference on World Wide Web. WWW '12, ACM (2012) 689–698
154. McPherson, M., Smith-Lovin, L., Cook, J.: Birds of a feather: Homophily in social networks. Annual Review of Sociology **27** (2001) 415–444
155. Macskassy, S., Provost, F.: Classification in networked data: A toolkit and a univariate case study. Journal of Machine Learning Research **8** (2007) 935–983
156. Mislove, A., Viswanath, B., Gummadi, K.P., Druschel, P.: You are who you know: inferring user profiles in online social networks. In: Proceedings of the third ACM

international conference on Web search and data mining. WSDM '10, ACM (2010) 251–260
157. Heatherly, R., Kantarcioglu, M., Thuraisingham, B.: Preventing private information inference attacks on social networks. Knowledge and Data Engineering, IEEE Transactions on **PP**(99) (2012) 1
158. Baatarjav, E.A., Phithakkitnukoon, S., Dantu, R.: Group recommendation system for facebook. In Meersman, R., Tari, Z., Herrero, P., eds.: On the Move to Meaningful Internet Systems: OTM 2008 Workshops. Volume 5333 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2008) 211–219
159. Zheleva, E., Getoor, L.: To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles. In: Proceedings of the 18th international conference on World Wide Web. WWW '09, ACM (2009) 531–540
160. Chaabane, A., Acs, G., Kaafar, M.A.: You Are What You Like! Information Leakage Through Users' Interests. In: Proc. Annual Network and Distributed System Security Symposium (NDSS). (2012)
161. Li, N., Li, T., Venkatasubramanian, S.: t-closeness: Privacy beyond k-anonymity and l-diversity. In: Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on. (april 2007) 106 –115
162. Machanavajjhala, A., Korolova, A., Sarma, A.D.: Personalized social recommendations: accurate or private. Proc. VLDB Endow. **4**(7) (April 2011) 440–450
163. Ying, X., Wu, X.: On link privacy in randomizing social networks. Knowledge and Information Systems **28** (2011) 645–663
164. Heatherly, R., Kantarcioglu, M., Thuraisingham, B.: Social network classification incorporating link type values. In: Intelligence and Security Informatics, 2009. ISI '09. IEEE International Conference on. (june 2009) 19 –24
165. Donath, J., Boyd, D.: Public displays of connection. BT Technology Journal **22**(4) (2004) 71–82
166. Gayo-Avello, D.: All liaisons are dangerous when all your friends are known to us. HT 2011 - Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia (2011) 171–180
167. Kubica, J., Moore, A., Schneider, J., Yang, Y.: Stochastic link and group detection. Proceedings of the National Conference on Artificial Intelligence (2002) 798–804
168. Zheleva, E., Sharara, H., Getoor, L.: Co-evolution of social and affiliation networks. In: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. KDD '09, ACM (2009) 1007–1016
169. Backstrom, L., Leskovec, J.: Supervised random walks: predicting and recommending links in social networks. In: Proceedings of the fourth ACM international conference on Web search and data mining. WSDM '11, ACM (2011) 635–644
170. Hogg, T., Wilkinson, D., Szabo, G., Brzozowski, M.: Multiple relationship types in online communities and social networks. In: Proceedings of the AAAI Symposium on Social Information Processing, AAAI (2008) 30–35
171. Hart, J., Ridley, C., Taher, F., Sas, C., Dix, A.: Exploring the facebook experience: a new approach to usability. In: Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges. NordiCHI '08, ACM (2008) 471–474
172. Krishnamurthy, B., Wills, C.E.: Characterizing privacy in online social networks. In: Proceedings of the first workshop on Online social networks. WOSN '08, ACM (2008) 37–42
173. Thelwall, M.: Social network sites: Users and uses. Advances in Computers **76** (2009) 19–73

174. Clemons, E.: The complex problem of monetizing virtual electronic social networks. Decision Support Systems **48**(1) (2009) 46–56
175. Boongoen, T., Shen, Q., Price, C.: Disclosing false identity through hybrid link analysis. Artificial Intelligence and Law **18**(1) (2010) 77–102
176. Enamul Kabir, M., Wang, H., Bertino, E.: A conditional purpose-based access control model with dynamic roles. Expert Systems with Applications **38**(3) (2011) 1482–1489
177. McNealy, J.: The privacy implications of digital preservation: Social media archives and the social networks theory of privacy. Elon Law Review **3** (2011) 133
178. Fisher, B., Turner, K., Morling, P.: Defining and classifying ecosystem services for decision making. Ecological Economics **68**(3) (2009) 643 – 653