12 Privacy in Social Networks

Traian Marius Truta, Michail Tsikerdekis, and Sherali Zeadally

1 Introduction

1.1 A Brief History

The worldwide Web has radically changed the way we communicate and interact with each other and how we manage our privacy. A good example of this is the ability to take photos that automatically include geographical information (often referred to as geotagging) and share them with a circle of friends. Traditional photos did not contain any geographical information and so questions that usually followed went along the lines of "where was this taken?" Such sentences are becoming obsolete and this is just one of the myriads of changes in our 21st-century digital lives. Of course change may not always be for the better. In the past decade we have seen cases where social media made news as the dangers of exposing one's private life where made apparent. Employees have been stalked online by

employers [62] and teenagers have been deceived by predators [6]. It seems that we are not yet fully familiar with this new world that came into our lives, or are we familiar?

T. M. Truta

Department of Computer Science, College of Informatics, Northern Kentucky University

Highland Heights, KY 41099, USA

e-mail: trutat1@nku.edu

M. Tsikerdekis and S. Zeadally

College of Communication and Information, University of Kentucky

Lexington, KY 40516, USA

e-mail: {tsikerdekis, szeadally}@uky.edu

Long before the advent of the Web, in the early 1990s there was a world of social media used in organizations to enhance collaboration [31, 32]. The motivation behind social media at that time stemmed from the need to collectively create and disseminate information and while early computer interfaces provided limited richness in people's communication, they were still effective enough to be adopted by organizations at the time. Bulletin board systems have been around since 1978 and have been used by people to make announcements, inform friends about meetings, and share other information through postings [65].

1.2 From Web 1.0 to Web 2.0

The revolutionary moment in history came with the advent of the Web or Web 1.0 in 1993 when it was released to the world [42]. While early web interface (e.g., gopher) provided the ability to view and edit pages (as it was the need of the early physicists at CERN (The European Organization for Nuclear Research) who needed to update and exchange results among them), it was, however, static and featured (technically) non-editable pages to individuals other than the owner of a server hosting those files. In fact, the Web remained this way for a while with people in 1999 describing web pages as "static screenfuls" [21]. There were various limitations as to the interactions provided by that early Web and so people who sought interactivity and exchange of content used software tools such as Internet Relay Chats (IRC) and MUD games [58]. Another prominent feature of Web 1.0 was the clear distinction between the user and the webmaster (the owner of a website). One-way communication between who contributes the content and

to whom it is being delivered could clearly be identified. All of this was bound to change the moment new technologies allowed for an advanced level of interactivity online. Adoption of new technologies seems to be dependent on the age with the younger population being more receptive to new technologies [57]. During the period 1995–2000, we saw an under-representation for the older age groups [48] and the adoption of new technologies was becoming more ubiquitous. Bernal [5] has been one of the few people to articulate the shift between Web 1.0 and 2.0. He argued that while the focus of Web 1.0 was on delivering products, the focus for Web 2.0 has been toward the delivery of services and increasing interactivity among users. Bidirectional interaction was quickly achieved by combining and ensuring compatibility among multiple technologies along with expanding the processing and scalability capabilities of databases and web programming languages. Additional service-oriented architectures helped to promote these services further. There was tremendous potential for many user-driven businesses to thrive under a Web 2.0 model [70] but many have also advised caution and suggested that this change may not ensure commercial success for all businesses [38]. Today, many enterprises are enjoying the benefits of Web 2.0 technologies with the majority of top executives favoring such strategies [60]. Web 2.0 technologies provide flexible design and rich and responsive user interfaces. They allow for collaborative creation of content, developing new application and services that communicate across different platforms, and establishing social networks of people with common interests, as well as supporting collaboration and collective intelligence [60]. It is worth pointing out that people were collaborating online and

forming communities well before Web 2.0 [39]. Howard [39] argued that the creation of online communities and collaboration could also happen with software (desktop applications or video games) that is not Web based. Gradually, a trend started appearing for Internet software that was providing more social tools to users. This is not surprising if one considers that users value personal interaction with the software as well as social interaction with other people [16]. The freedom provided by interactive social tools that allowed not only for two-way communication between users but also user contributions to content enabled Internet users to explore social interactions like never before. Networked communications have evolved to accommodate the needs of humans as social beings [40]. The idea of social media came to life.

2 Social Media

Social media and Web 2.0 are not the same. Social media refer to Internet-based applications that build on the foundations of Web 2.0 and allow for the creation and exchange of user-generated content [43]. Under the large umbrella of social media one can find applications which include, blogs, collaborative project (e.g., Wikipedia), social networking sites (e.g., Facebook), content communities (e.g., YouTube), virtual social worlds (e.g., Second Life), virtual game worlds (e.g., World of Warcraft), and micro-blogging (e.g., Twitter) [43, 44]. Social networks have had a great impact on our society and they are the most representative type of social media for their use of Web 2.0 technologies.

2.1 Social Networks

Social networks have gained a lot of interest and popularity over the last decade. Kaplan and Haenlein [430] defined them as applications that enable users to create personal profiles, invite friends to connect with them, and to have access to other people's profiles. These profiles can include various types of information such as photos, video, audio files, and even blogs. The basic ingredients of a social network are to allow for the construction of public or semi-public profile, to articulate a list of users that individuals share a connection with, and to view and share that list with others within a system [10]. There is also a distinction between a social network site and a social networking site. According to Boyd and Ellison [10], networking implies relationship initiation often between strangers. However, lines have been blurred with today's social networking services offering both networking with existing relationships as well as initiating new with strangers. Henceforth, we assume that by social networks we mean applications both for network as well as networking in terms of the goals of a social media service.

2.2 Social Networking Sites

The first social networking site according to many was SixDegrees.com and was launched in 1997 [10]. It was the first website to combine features that allowed profile creation, forming friend lists, and sharing those lists with others. The website managed to attract 3.5 million users until it finally closed down in 2000 after being bought off for $125 million [51]. Subsequently, several other services such as LiveJournal.com started offering social networking features [10], but it was re-

ally later on in 2003 when modern social networking sites were launched with the primary goal of providing a digital representation of user networks, initiating and managing relationships.

2.2.1 LinkedIn

Linkedin.com was launched in 2003 with the intent to connect professionals with their networks. In January 2009 the network had 32 million members and in March 2011 it had 100 million members.[1] At the time of writing, the website has 225 million users.[2] LinkedIn allows individuals to create professional networks, to view how they are linked with other members, and view what their degree of separation is (how many connections apart they have) from a target member [49]. This means that an individual's social network becomes tangible. As such, social capital has ceased to be an abstract concept but has become a visible structure that an individual can keep expanding and restructure.

2.2.2 Friendster

Another website that was launched in 2002 did not share the same success that LinkedIn did. Friendster is recognized as one of the best examples of early popular social networks [10]. The website started off as a dating website but encouraged users to join even if they were not looking for dates [8]. The idea behind Friendster was that friends of friends are good candidates for dates. The decision was made to arbitrarily allow people to connect with others as far as four degrees

[1] http://techcrunch.com/2013/01/09/linkedin-hits-200-million-users-worldwide-adding-new-users-at-rate-of-two-per-second/

[2] http://www.linkedin.com/about-us

(connections between individuals) away in their network. Any individuals beyond four degrees from an individual could not be reached; a choice that is restrictive for a community according to the theory of six degrees of separation [76]. The website was launched in 2002 and by mid-August it had 1.5 million registered users [8]. Boyd [9] was one of the first researchers to study the popular website and suggested that the human–computer interaction community should consider the evolution of social community along with the underlying technology. Her arguments made an accurate prediction of the technical and social difficulties that the website later experienced. Servers frequently failed because they could not sustain the increased traffic as premature web software of the time was not designed to handle the amount of interactive actions of millions of users. In turn, users became frustrated, leading to some of them switching their email addresses to Friendster's email service. Additional social issues (such as the influx of new users who were unfamiliar with community norms) also led to the decline of Friendster's online community. The balance of current social groups was shaken due to the influx of new users and users who wanted to connect with others beyond the four-degree limit [8]. The so-called fakester account was an early version of developing pages of special interests so that people can find others with common interests (e.g., fans of *Star Wars* movies). Many of these accounts had thousands of friends, which created computational loads for the ill-equipped servers at the time. The decision was made by Friendster to delete all of these accounts to resolve website issues.This resulted in a rejection of the website in the United States by early adopters due to several issues such as social collisions (e.g., employers being able

to monitor their employees' work activities) and a loss of trust between users and the site as a result of the deletion of these accounts [25]. Many of these actions violated the hierarchy of needs for online users, which arguably if used could have put user needs first [50]. The website has made a comeback in recent years and in October 2008, according to a press release, it reached 85 million members worldwide.[3]

2.2.3 Myspace

Friendster was followed by Myspace, another popular social networking site that was launched in August 2003. Myspace grew rapidly as Friendster's popularity declined, because some of their adopters saw it as a safe haven to express their interests (something that was limited in Friendster due to its four-degree policy) [10]. Significant attention was given to bands and music, which helped to increase the number of users. Myspace expanded its features based on user demand and allowed for page personalization (e.g., adding HTML to alter the layout), which boosted its popularity further. Myspace also focused on developing policies to allow teenagers to join the service, which further increased its user base. At its peak, in 2008, the website had 75.9 million users[4] before the service started declining because of safety issues that plagued the service [10]. In June 2011, the service was down to 33 million users although after a recent redesign it has been picking up traffic once again.

[3] http://web.archive.org/web/20100522004359/http://www.friendster.com/info/presscenter.php?A=pr48

[4] http://mediadecoder.blogs.nytimes.com/2012/02/12/myspace-to-announce-one-million-new-users/?_r=0

2.2.4 Facebook

One social networking service that perhaps gained from all the predecessor social networks that rose and fell was Facebook. It is the most popular social networking site currently and the longest to maintain such a title. The service has experienced a skyrocketing growth by designing its website to provide the best features by addressing several of the deficiencies of previous social networking services. Launched in 2004, Facebook has seen a dramatic increase of its user base worldwide (Fig. 12.1). In September 2013, the website had 1.19 billion users monthly with average daily unique users at 727 million.[5] Approximately 80% of its daily user base is outside of the United States and Canada, with some countries reaching high penetration levels among their Internet users (higher than 90%).[6]

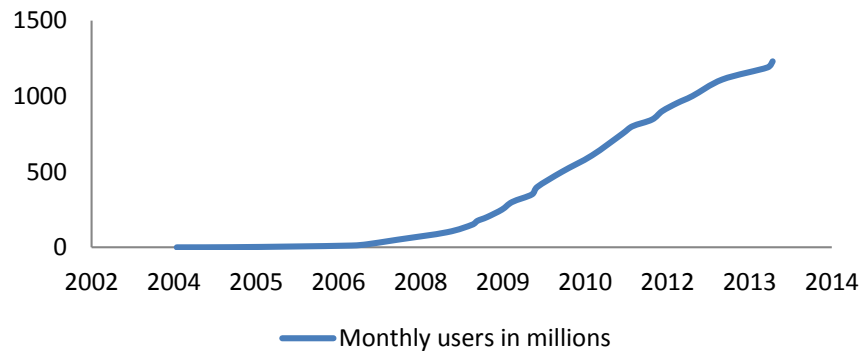


Fig. 12.1 Approximate growth of monthly users in millions for Facebook

(Source: https://newsroom.fb.com/)

[5] http://newsroom.fb.com/Key-Facts

[6] http://www.internetworldstats.com/facebook.htm

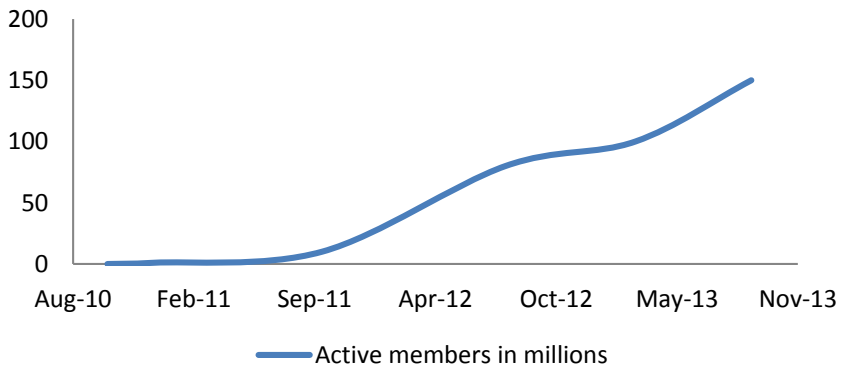
2.2.5 Mobile social networking

Social network usage has increased by 64% since 2005 [11]. Currently, Facebook and Twitter (a micro-blogging service) have reached 82% of the world's Internet users [69]. In the last few recent years a dramatic shift has been observed in people accessing the Internet via mobile devices leading to the emergence of mobile social networking. Mobile social networking implies social networking services, which include social structures with entities (individuals or organizations) connected through various types of interdependency (e.g., common interest, friendship, etc.), that are used by individuals through their mobile devices [41]. Jabeur et al. [41] attribute the rise in popularity in the enabling of new ways for social interaction and collaboration by taking advantage of location-based services and data-sharing services (e.g., photos) provided by mobile devices in an immediate way. Mobile social networking services can be divided into two types, those with native support only for mobile devices (e.g., Instagram) and those offering mobile as well as web access to their services (e.g., Facebook).

Historically, early mobile social networking has been observed since 1999 [52]. These applications came usually pre-installed in mobile devices and some followed a subscription-based model. They are similar to primitive versions of early social networks with the ability to broadcast messages to many people at once, but focus less on profile creation and management. During the early 2000s, a transition was observed with the release of early wireless application protocols (WAP) third generation (3G) technologies when applications started being released with social networking features incorporated in them. These were still developed and

maintained by the manufacturer, or in close association with the manufacturer or carrier of mobile services. By the late 2000s, applications developed by third parties (e.g., independent developers) were able to be installed in mobile devices, which radically altered the range of applications available for consumers.

One of the most popular examples of early native mobile social networking applications was Instagram, which was launched in October 2010. The application provided a photo and video sharing social networking service to mobile users in collaboration with other social networking services (through websites and mobile portals). The service was released for free through Apple's App Store and Google Play, which helped to increase its popularity. By April 2012 it had 100 million active users[7] when it was sold to Facebook for $1 billion.[8] Figure 12.2 depicts the growth of the service.



[7] http://instagram.com/press/

[8] http://abcnews.go.com/blogs/technology/2012/04/facebook-buys-instagram-for-1-billion/

Fig. 12.2 Growth of active members in millions for Instagram

(Source: http://instagram.com/press/).

Many social networking sites also expanded their access to mobile devices. Facebook started offering mobile access to iPhone users in August 2007 and almost a year later it reached 1.5 million regular users. In 2008, a Facebook mobile application was offered to iPhone users. As of December 2013, 945 million users access Facebook monthly through mobile devices (approximately 77% of its total monthly users).[9]

The increase in usage of mobile social networks has led to the emergence of geosocial networking. This is social networking that includes geographic services and features such as geocoding and geotagging, which alter the social dynamics of a mobile social networking service (e.g., recommendation systems that can help with attendance at events in close proximity based on past movement patterns and location history) [64]. For web-based social networks, a user's location is attached to content using their internet protocol (IP) address (which is tracked to an approximate position at city or area level) or wireless hotspot trilateration (which uses multiple wireless hotspots to determine the relative location of a user). For mobile social networks, cell phone tracking and Global Positioning System (GPS)-enabled services can be used to attach geographical information to content.

[9] https://newsroom.fb.com/key-Facts

2.3 Impact of Social Networking on Society

The success of social networking sites can be attributed to their ability to satisfy social needs (e.g., the need to communicate with others and be a part of a social group) that online users have. Social networking sites have become an extension of an individual's real life, containing a detailed documentation of a person's social network along with aspects such as their experiences, thoughts, beliefs, and preferences. Social networking sites are helpful for people with low self-esteem and low life satisfaction and provide a tremendous advantage for managing social capital [25]. Social capital, defined loosely as the value of social relations that helps provide benefits to individuals or groups [17], became the term to define the well-being of groups and society. As the number of social networking users increases, a higher number of online relationships are expected to form, and, as a result, people connected to others are likely to receive more positive feedback from these relationships [77]. Positive feedback received by users' social networks enhances their social self-esteem as well as their well-being. People using social networking sites tend to have more virtual friends than real-life friends [79]. Corporations also exploit the benefits of using social networks for supporting brand promotion and marketing campaigns [12, 17]. Social networks can also be profitable business models [59].

Social networks have also been affected by various issues. One example is the differential adoption due to the digital inequality [35]. This digital divide has economic, sociological, and political drivers that affect not just the adoption of social

networking sites but also the adoption of the Internet [30]. For users who end up using social networks, one of the most popular issues relates to privacy [3, 53, 68].

3 Privacy Issues in Social Networks

As discussed in the introduction to this chapter, there are a large number of privacy concerns in the field of social networks. These concerns have greatly increased in the past years due to the advent of online social networks. Facebook, LinkedIn, and Twitter are already well-known social networks that have a large audience in all age groups. Recently more trendy social sites such as Pinterest, Instagram, Vine, Tumblr, WhatsApp, and Snapchat are being preferred by the younger audience [63]. The amount of data that those social sites gather from their users is continually increasing and these data are very valuable for marketing, research, and various other purposes. At the same time, the data usually contain a significant amount of sensitive information, which should be protected against unauthorized disclosure. It is safe to say that any collection and storage of individual data regardless of intent, can lead to privacy implications that would not have existed otherwise [66]. One example of such a situation was in 2006 when, to stimulate research on real Internet data, AOL made available over 20 million search queries from over 650,000 users. Although the data was de-identified (in a poor way), individuals that conducted specific searches were identified in the data. The main reason why this was possible was that many users searched for their city, neighborhood, and even their first and/or last name. *The New York Times* published a

story about one such re-identified individual, Thelma Arnold, from Liburn, Georgia in the United States, who was discovered through her queries terms [2]. Luckily, no significant harm was reported for any individual from the released data. However, the researcher responsible for de-anonymizing and releasing the data was dismissed and the AOL chief technology officer resigned.

For social network data, privacy can be seen from different angles. Imagine an online social network site (such as Facebook, Orkut, etc.). These sites gather data from a large number of users, and that data is published to other users based on privacy controls of the user that owns the data. For instance, Facebook has a series of privacy settings that allows a user to choose what to share and with whom. These controls go beyond these basic features, and a user can create various levels of friends, review any information that others post about them before it is posted, and so on. What is important to note at this point is that this view of privacy is user-centric or local. This type of privacy is commonly called *social privacy* [66]. A second view of privacy is when we look at the whole social network data. Any social network site will gather data and use this data for other purposes as specified in their data use policy. For instance, Facebook has a very detailed data use policy in which they describe how they use the information received from their users. Of particular interest for privacy is how this information is shared to other parties (companies):

"Your trust is important to us, which is why we don't share information we receive about you with others unless we have:

- received your permission;

- given you notice, such as by telling you about it in this policy; or
- removed your name and any other personally identifying information from it." [26]

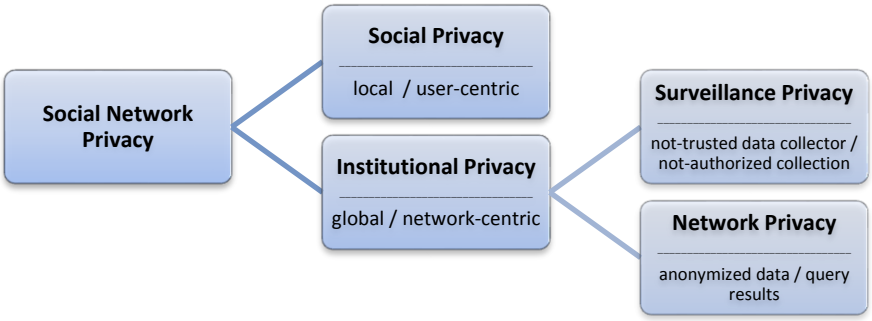


Fig. 12.3. Social network privacy types.

As stated above, the social network data is de-anonymized prior to being shared to other companies. However, as seen from the AOL case, the de-anonymization process may not be fully successful and the privacy of certain individuals may still be at risk. This view of privacy is network-centric or global and it is commonly called *institutional privacy* [66]. The institutional privacy can also be seen from two distinct angles. First, the social network site, as the data collector (many times this is referred to in the literature as data owner, we chose to use this term since in many data use policies, such as Facebook's, the data owner is considered the user that provided the data), has unlimited access to all collected data, thus, protecting privacy from the data collector is an Herculean task. In general this situation is not considered a privacy concern because the data collector is trusted with the data di-

rectly by the user. The user has the option of not participating in that social network site and he or she remains unknown to the data collector. This is more difficult than it appears because in many cases the data is not voluntarily provided to a data collector. An example of such a situation is the data collection practices of NSA as revealed by the whistle-blower, Edward Snowden [24]. This type of privacy, when the social network data collector is not trusted or the data is gathered without the knowledge of the user, is known as *surveillance privacy* [66]. The second view of institutional privacy is when the social network data is shared by a trusted data collector to third parties. Due in major part to AOL anonymization failure, there are no recent attempts to publically provide anonymized data to researchers; however, this sharing of collected data happens when there is a significant benefit for the social network site. The data is anonymized (this is most likely specified in the data use policy, for instance, Facebook will anonymize their data before sharing it with others) and shared with companies that are in general trusted by the original data collector. However, the anonymization process must aim to protect the individual data from disclosure in case attempts to re-identification occur. In the context of social network data, we call this type of privacy *network privacy*. A variant to this scenario is when the data is not shared with other parties, but the data collector shares the result of various queries with third parties. While this approach seems to better protect the individual's privacy it still may lead to privacy breaches and it requires the data collector to be able to process the queries requested by other parties, anonymize the query result, and provide these results to

requestors. We include this scenario in the context of network privacy. Figure 12.3 illustrates these privacy types.

Table 12.1 Social network privacy concerns

Social Privacy	*User awareness*
	Privacy controls complexity
	Privacy controls changes
	Privacy controls conflicts
Surveillance Privacy	*Not-trusted social network provider*
	Data collected without user permission
	No oblivion
Network Privacy	*Data collected for profit*
	Lack of proper anonymization
	Increase sharing of collected data

We will present briefly the main privacy concerns related to each type of social network privacy (see Table 12.1 for a summary). A solution for each such problem is presented in the next section. For *social privacy*, the main concern is whether or not the user understands the privacy risks he or she is taking when sharing information on a social network (*user awareness*). As recent as 2012, ap-

proximately 8% of US Facebook users had never heard about Facebook privacy tools. What is more alarming is that even people that are aware of privacy risks do not take appropriate steps to protect their privacy. For instance 28% of US Facebook users share their wall posts to a wider audience then their friends [18]. The positive news is that users have become more aware of their privacy. In a study that used public profiles from New York City, 52.6% of the users hid their friends list from their public profile as of June 2011, whereas in March 2010, only a little bit more than a year earlier, 17.2% of the users hid their friends list [19]. Related to the user awareness with respect to privacy, difficulty in setting privacy controls makes the users prone to giving up in selecting an appropriate privacy policy (*privacy controls complexity*). For example, Facebook privacy controls are spread in at least six different tabs: Privacy, Timeline and Tagging, Blocking, Followers, Apps, and Ads. An example of such a tab is shown in Fig. 12.4. To add to this complexity, the privacy controls are not easily accessible from the data use policy, and when there explanation is not clear or even provided [27].

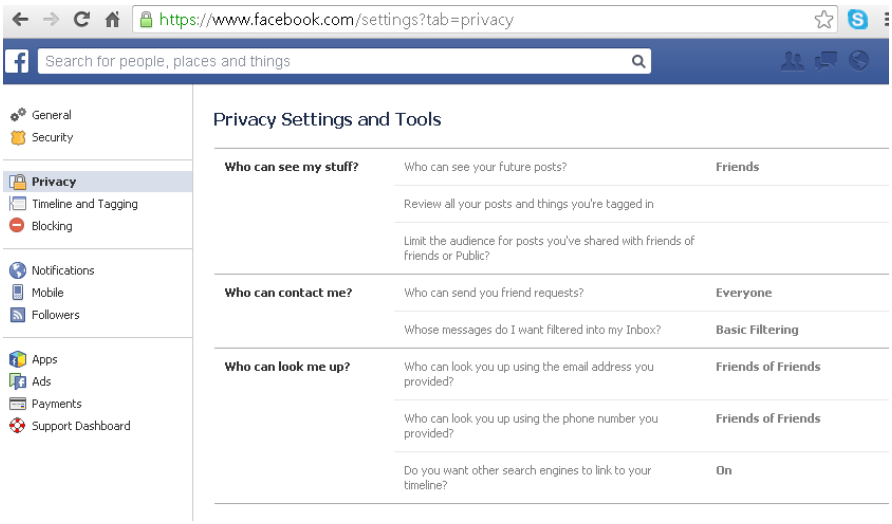


Fig. 12.4 Facebook privacy control — privacy tab.

Moreover, privacy controls may change and this can contribute to reducing the privacy (*privacy controls changes*). Again, we use Facebook as an example. As recent as late 2012, Facebook made significant changes to their privacy controls and policies. While these changes simplified the privacy control and policies, they create some additional privacy concerns. For instance, Facebook decided to remove the privacy setting that let users hide their Timeline from people who search for it [34]. In addition, some privacy shortcuts were disabled and made available only from the main privacy page. An example of such a privacy shortcut is the pop-up on the top of the News Feed that answered questions such as "Who can see my stuff?" [34]. Also, the data use policy does not offer direct links to privacy

controls. To add to that, the Facebook privacy policy changed to allow more sharing of data to third-party companies. The new policy states:

> You give us permission to use your name, and profile picture, content, and information in connection with commercial, sponsored, or related content (such as a brand you like) served or enhanced by us.

While the old policy was more user-friendly:

> You can use your privacy settings to limit how your name and profile picture may be associated with commercial, sponsored, or related content (such as a brand you like) served or enhanced by us. [27]

This new policy is more related to network privacy and it shows that some of the privacy concerns are applicable to more than one privacy type.

In some cases, the privacy controls may have conflicts and, when two controls specify the privacy setting for the same data item, it is difficult to know which privacy control takes precedence (*privacy controls conflicts*). Privacy policy conflicts exist in many common social networks such as Facebook, MySpace, Orkut, Twitter, and Google+ [80]. For example, in Facebook, a user may choose to have his or her friends' list private. However, if some of that user's friends keep their corresponding friends list public, some friendship relations can be inferred by an authorized user. This type of conflict is common to other privacy settings as well [80].

With respect to *surveillance privacy*, an important concern is that the initially trusted social network becomes non-trusted (*not trusted social network provider*). Also, there are organizations that have the capability of collecting data without user approval and can use this data for their own purposes (*data collected without user permission*). In addition to these concerns, the fact that any published data may stay published or stored forever may increase the possibility of surveillance

and constitute an important privacy concern. It is very difficult to enforce the right to be forgotten, also known as oblivion, on social networks (*no oblivion*). Different countries have opposing views with respect to oblivion and their regulations are contradictory to each other. For instance, in France, the law recognizes the right of oblivion, a convicted criminal can object to the publication of his criminal record after he or she has satisfied their punishment. In the United States, publication of criminal records is protected by the First Amendment.

Network privacy concerns are less known to the general user of a social network than the social and surveillance privacy concerns, but they are very important in any discussion of social network privacy. The main reason a social network site gathers user data is to be able to monetize that data. Gathering more personal data, which can be successfully analyzed, mined, and consequently used for target advertisement, is the main goal of a social network company. This ever increasing amount of personal data creates more and more potential privacy violations (*data collected for profit*). In the past few years, Facebook users disclosed less information publically, which shows increase in user awareness of social privacy concerns. However, during the same time, the average Facebook user seems willing to disclose more and more information privately to his or her friends. This contributes to more data collected by Facebook and third party apps, and this data can be used for advertisement or other purposes directly by the data collectors [74]. The collected data are usually released to other companies in an anonymized form; however, since the anonymization methods are not public, it is not clear if the anonymized data are able to avoid re-identification of individuals (*lack of*

proper anonymization). For instance, Facebook can share user data if they "remove your name and any other personally identifying information from it." Currently, more and more companies are specialized in Big Data and data analytics. Developing efficient methods to analyze large amount of data will contribute to a need for social network data. A social network site will benefit from selling their anonymized data to such data analytics companies and potential privacy violations will increase (*increase sharing of collected data*).

The above classification is not completely disjointed; some of the privacy concerns are true for more than one privacy type. For instance, *user awareness* is also important for surveillance privacy and network privacy, and *no oblivion* privacy concerns exist in network privacy as well.

Section 4 will provide existing privacy solutions to the above concerns with a focus on technical solutions.

4 Privacy Solutions for Social Networks

Since there are many privacy concerns regarding social network data, there is not an easy solution to these problems. Moreover, to protect privacy of individuals the privacy solutions must be supported and provided by legislators, social network sites (social network service providers), and social networks users [67]. All these three entities have the ability to enhance the privacy protection for each type of entity. Figure 12.5 captures this interaction. Social network privacy is divided between social privacy, surveillance privacy, and network privacy (institutional pri-

vacy is not shown). The legislators, social network sites, and their users can provide privacy solutions for each type of privacy.

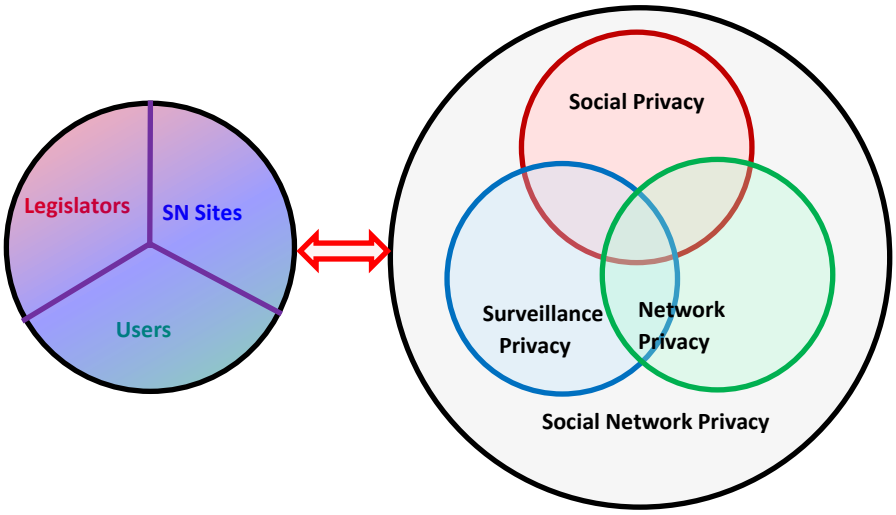


Fig. 12.5 Social networks privacy — a common effort.

For *social privacy*, the legislators can require that social network sites have a privacy policy and a set of privacy controls that is appropriate for the type of data the site collects. The legislators can also require that the social network sites have a good education system of their users and the privacy implications of their data are disseminated to all their users. The social network sites also provide important solutions for social privacy concerns. Privacy friendly default settings, easy to use privacy controls that change infrequently or not at all, allowing creation of pseudonymous profiles as an option, and avoiding privacy conflicts are some of the solutions a social network site can employ to protect the privacy of their users. Last but not least, the users must be educated about the privacy implication of sharing

their data. In the context of social privacy, the users should make sure who their friends are, and they should use appropriate privacy controls for the data they share. It is important to respect the privacy of others as well, and to also guard the privacy of one's children [67].

We provide an example regarding privacy policy conflict and we discuss how this problem can be solved.

Fig. 12.6 Allow-take-precedence privacy policy

In Fig. 12.6, the Celebrity user chose to make her list of friends private. Some of her friends (Friend 1 and Friend 2 are depicted) chose to make their list of friends public. Due to their choice, the corresponding friendship relationships are public and this violates the choice of the Celebrity user. This privacy policy conflict, known as *allow-take-precedence policy*, is widely used in existing social networks such as Facebook and Orkut [80]. Solutions proposed for this privacy violation include [80]:

- **Redesign of privacy policy.** This is extremely difficult if the users can choose their own privacy policies. While it is easy to employ, it will set standard privacy policy for all users that can be viewed as either too restrictive or too permissive.
- **Deny-take-precedence policy.** The social network site may deny the Friend 1 and Friend 2 users the ability to publish their friendship relation with Celebrity user due to Celebrity user settings. This approach is known as deny-takes-precedence. Since it is based on both users' preferences, it requires more processing from the social network site software and it is not currently employed. This approach will give preference to privacy when there is a privacy policies conflict between users.
- **Avoid using bi-directional friendship relations.** This is possible in social networks that allow relations of type followers and following. In this case each user may choose their own preference for their corresponding lists. Still an adversary may infer entries of a private list from public lists of the victim's friends (followers or followings), and these solutions still have the original problem although in a limited scope.
- **Privacy policy negotiations.** In this scenario, privacy policies are dynamically updated based on given requirements of utility and privacy. Such policy negotiations are still in an early development stage and it is not clear how well they can satisfy all users. As an example, in a game theoretic approach used for those negotiations, users cannot protect their information if others sharing the information request to make it available [73].

For *surveillance privacy*, the most obvious solution is to avoid posting any sensitive information on online social networks. While this is an easy solution, it is difficult to enforce considering how pervasive the social networks are today. In this type of privacy, the social network site is not trusted and thus the private information should not be provided in clear form. The basic solution for enforcing this is the usage of cryptographic methods. There are several applications that use encryption to protect users' information on the social network sites. Some of them are listed below:

- **FlyByNight.** This application is implemented for Facebook and encrypts the user data before being stored on Facebook. Unfortunately, FlyByNight relies on Facebook servers for key management, so it fails to protect against the surveillance of the social network provider []55.
- **NOYB (none of your business).** NOYB is also used on Facebook and it uses encryption to protect personal details of users. It protects against the surveillance of the social network provider (Facebook in this case) but it is applicable only to specific attribute data from the user profiles and it does not allow encryption of free text [29].
- **FaceCloack.** This application is a Firefox browser extension that uses a symmetric key to encrypt user personal information in Facebook. This method requires the use of dedicated FaceCloack servers that store part of the user profile in an encrypted form [56].

- **Scramble!** This application is designed as independent from a specific social network platform. The content is also encrypted prior to being shared in the OSN, and only friends can decrypt it [4].

Other solutions regarding surveillance privacy include implementation of a social network site as a distributed site, use of fake traffic to obscure user activity, and use of anonymous communication network such as Tor [20].

The main solution for *network privacy* is network anonymization. To define a network anonymization model it is important to understand what constitutes a privacy violation for a social network. A *privacy violation* (or breach) occurs when sensitive information about an individual is disclosed by an adversary. In the context of social networks the most common types of privacy violations are: *identity disclosure*, *attribute disclosure*, and *link disclosure* [81].

Identity disclosure refers to the correct re-identification of a node (such as a person or an institution) in an anonymized social network when the adversary uses the anonymized network and other available information about individuals from the network.

Attribute disclosure refers to an adversary finding out something new about the target individual, but in this case the adversary may not know which node in the network the individual represents.

Link disclosure occurs when an adversary discloses the existence of a sensitive relationship between two individuals from the social network. This type of disclosure assumes that some relationships are sensitive and their privacy must be protected.

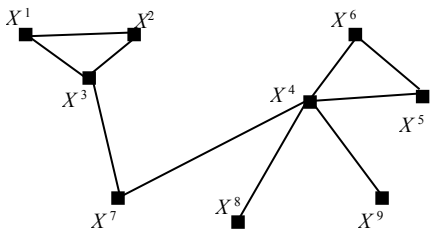
In order to anonymize a social network it is also important to understand what types of data are sensitive and what types of data might be known from other sources. These assumptions lead to various *social networks models*. We present below an example of such a model.

We model a social network as a simple undirected graph $G = (N, E)$, where N is the set of nodes and $E \subseteq N \times N$ is the set of edges. Each node represents an individual entity. Each edge represents a relationship between two entities.

The set of nodes, N, is described by a set of attributes that are classified into the following three categories: *identifier* attributes such as *Name* and *SSN* that can be used to identify an entity, *quasi-identifier* attributes such as *zip code* and *sex* that may be known by an adversary, and *sensitive* attributes such as *diagnosis* and *income* that are assumed to be unknown to an adversary.

For simplicity, only binary relationships are allowed in our model. Moreover, all relationships are of the same type and, as a result, they are represented via unlabeled undirected edges. Also, this type of relationship is considered to be of the same nature as all the other "traditional" quasi-identifier attributes. In other words, the graph structure may be known to an intruder and used by matching it with known external structural information; therefore, serving in privacy attacks that might lead to identity and/or attribute disclosure. In this model, link disclosures are not a privacy concern. An example of a social network is shown in Fig. 12.7. *Age* and *zip* are quasi-identifier attributes and *disease* is a sensitive attribute. The identifier attributes are not shown.

In addition to the privacy concerns that must be understand and captured in an anonymized network, of similar importance is the utility of the data. The anonymized network, while protecting the individual's privacy must also preserve much information to maximize the utility of the social network. Since it is difficult to know how the network is used, defining utility is not a trivial problem. Early work in social network anonymization usees the total number of edge additions and deletions to measure the utility loss [54]. Newer approaches focus on preserving the topological features of a network such as centrality measures, degree distributions, and clustering coefficients [1].



Node	*Age*	*Zip*	*Disease*
X^1	25	41076	diabetes
X^2	25	41075	cancer
X^3	27	41076	flu
X^4	35	41099	cancer
X^5	38	48201	cancer
X^6	36	41075	flu
X^7	30	41099	flu
X^8	28	41099	diabetes
X^9	33	41075	diabetes

Fig. 12.7 A social network example

We present next some of the most common social network anonymization approaches.

The main two approaches to anonymizing social networks are categorized as follows [81]:

- **Edge modification.** These techniques propose edge deletion and additions to help in anonymizing the network. The network structure will be altered by these changes, and the goal is to minimize the number of edge modifications while the privacy requirements are met and the data utility is maximized. The most used anonymization approaches in this category are: *k-degree anonymity* [54], *k-neighborhood anonymity* [82], and *k-automorphism* [83]. These approaches will be briefly introduced in this section. The above models focus on avoiding node re-identification. Other approaches such as *k-isomorphism* [15] and *l-opacity* [61] focus on preventing link disclosure, in which the adversary learn about a sensitive relationship between individuals.
- **Clustering or network generalization.** This technique proposes publication of aggregate information about the network structure. In this way attacks based on network structure are made very ineffective; however, the utility of the network may be too low. We will present the *k-anonymous clustered social network* [13, 75] in this section.

Two other approaches are as follows:

- **Randomization.** This is a special case of anonymization via edge modifications. The graph structure is modified by deleting and adding edges at random such that the total number of edges is unchanged. Unfortunately, this approach is altering significantly the utility of the data [36].

- **Differential privacy.** In this approach individual nodes are protected under the definition of differential privacy [23]. Usually in this approach the network is not anonymized and it is kept by the data owner, only releases of network measures such as degree distribution are allowed [37]. This constraint makes the differential privacy approach less flexible than the other anonymization approaches mentioned above. However, very recent developments allow non-interactive network data publication while differential privacy property is satisfied [14]. A high-level discussion about *differential privacy in social network data* is included in this section.

The *K-degree anonymity* model assumes that the degree sequence of nodes in a social network is potentially available to an adversary and the anonymization aims to create groups of nodes with similar degree values. A network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ is k-degree anonymous if for every node $X \in \mathcal{N}$ there exist at least $k - 1$ other nodes that have the same degree as X. Liu and Terzi proposed an algorithm that creates a k-degree anonymous network and minimizes the number of edge deletions and additions [54]. In Fig. 12.8 we illustrate an example of a three-degree anonymous network. Notice that three new edges were added to the network (shown in bold) and one was deleted (dashed). In this example nodes X^1, X^3, and X^4 have the degree 4, and all other nodes have the degree 2.

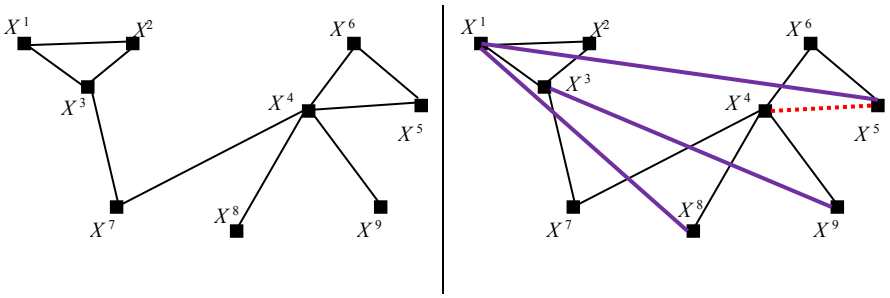


Fig. 12.8 A social network and a corresponding three-degree anonymous network

The *k-neighborhood anonymity* model assumes that adversary knows the immediate subgraph of the target node. The immediate subgraph contains all neighbors and relationships between neighbors. A node X is k-neighborhood anonymous if there exist at least k-1 other nodes such that the subgraph constructed by the immediate neighbors of each such node is isomorphic (has the same structure) to the subgraph constructed by the neighbors of X. By immediate neighbors we mean the nodes that are directly connected to the starting node. A graph satisfies k-neighborhood anonymity if all the nodes are k-neighborhood anonymous. There are heuristic algorithms that construct k-neighborhood anonymous networks. Such algorithms start by identifying all different neighborhoods and then it creates groups of identical neighborhoods of size k using edge additions and deletions [82]. In Fig. 12.9 we show a three-neighborhood anonymous network. Notice that three new edges were added to the network (shown in bold) and two were deleted (dashed). In this example nodes X^3, X^4, and X^7 have isomorphic immediate neighborhoods. All the remaining six nodes have also isomorphic neighborhoods.

K-automorphism anonymity assumes that the adversary can know any subgraph around a certain node. A network is *k*-automorphic if the view of the network from any node is identical with the view of the network from at least *k*-1 other nodes. The complete mathematical definition for *k*-automorphism and a heuristic algorithm is presented in [83]. Note that in Fig. 12.9, the anonymous network is also *k*-automorphic.

Based on the above definitions, it is easy to notice that any *k*-automorphic network is also *k*-neighborhood anonymous network, and any *k*-neighborhood anonymous network is also *k*-degree anonymous network.

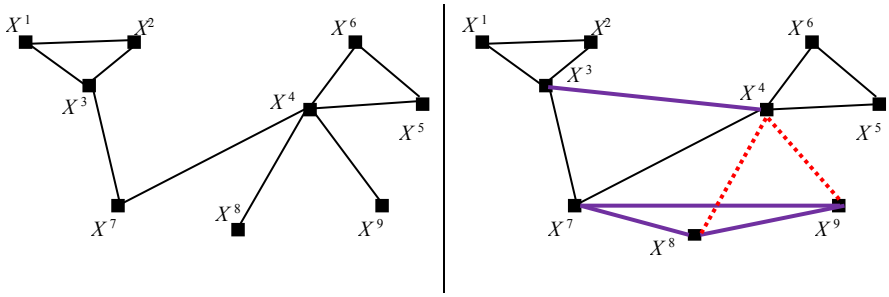


Fig. 12.9 A social network and a corresponding three-neighborhood anonymous network (which is also three-automorphic network)

A *k-anonymous clustered social network* uses a different approach. Based on a grouping strategy that tries to maximize an objective function, the nodes from a network are partitioned into pair-wise disjoint clusters. These clusters will then be

generalized to super-nodes, which may be connected by super-edges. The goal of this process is to make any two nodes coming from the same cluster indistinguishable based on their relationships. To achieve this objective, Campan and Truta developed intra-cluster and inter-cluster edge generalization techniques that were used for generating super-nodes and super-edges, and so generalizing the social network structure [13]. To satisfy the *k*-anonymous clustered model — derived from the well-known *k*-anonymity property for microdata — each cluster must have at least *k* nodes. The algorithm used in the anonymization process, called the SaNGreeA (Social Network Greedy Anonymization) algorithm, performs a greedy clustering processing of an initial social network in order to generate a *k*-anonymous clustered social network. In this algorithm the nodes that are more similar in terms of their neighborhood structure are clustered together using a greedy approach. To do so, a measure that quantifies the extent to which the neighborhoods of two nodes are similar to each other is used. Full descriptions of this measure and of the SaNGreeA algorithm are presented in [13]. Improving the SaNGreeA algorithm, Tassa and Cohen introduced a more efficient algorithm, namely sequential clustering algorithm, for creating *k*-anonymous clustered social network. Details about this new algorithm and a complete comparison in terms of both efficiency and utility with SaNGreeA can be found in [75]. Figure 12.10 shows two three-anonymous clustered networks.

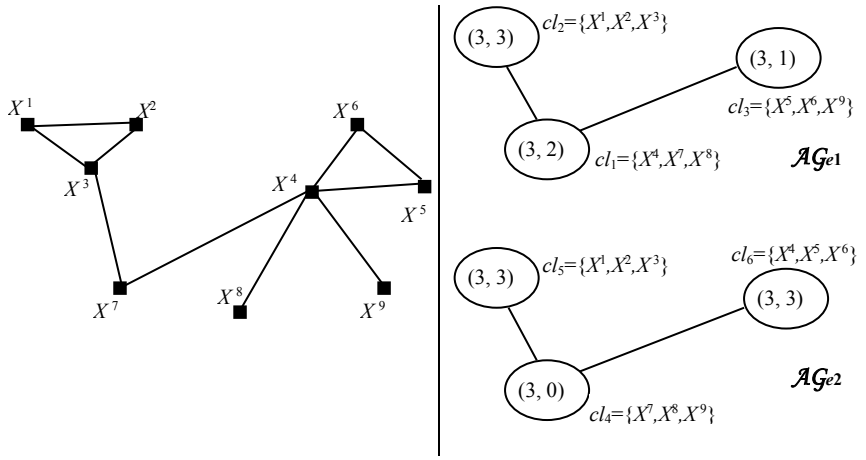


Fig. 12.10 A social network and two corresponding three-anonymous clustered social networks

Differential privacy in social networks is a new research direction that extends the differential privacy for tabular data to networks. Differential privacy is based on a mathematical guarantee of privacy which states that anything that is learnable from a table *T* can also be learned from a table *T'* which differs by only one record from table *T* [23]. Such a table *T'* is called a neighboring table for *T*. In case of networks, the notion of vicinity or neighboring can be defined in terms of both edges and nodes. Based on this, two models were created, edge differential privacy [33, 45, 46] that defines neighboring networks that differ by at most one edge, and nodes differential privacy [7, 47] in which neighboring networks differ by one vertex and its corresponding edges. Until 2014, all of this work was based on an interactive setting, which means that a trusted curator that has access to the original network will receive queries from non-trusted users and will apply a differentially private algorithm to provide the answer to users. Each user will have a pri-

vacy budget that can be exhausted if too many queries are sent to the curator. Recently, one practical solution for non-interactive network data publication was introduced in [14]. This solution, called *density-based exploration and reconstruction (DER)* creates a sanitized network G_s from the original network G that satisfies ε-differential privacy for the privacy budget ε. In addition to differential privacy requirements, this models aims to provide privacy guarantee even for correlated data (the original differential privacy model assumes independent data) if the amount of correlation can be measured. Full details regarding this approach can be found in [14].

5 Challenges and Opportunities in Social Networks Privacy

As already presented in the previous sections, there is not a universal solution to social network privacy, and there are many reasons for this.

While in other domains such as healthcare or financial sectors there are privacy regulations that define an expectation of privacy, in the social networks privacy is not as well defined, being interpreted differently by various users and social network sites owners. *Common interpretations of privacy in social networks as well as regulations that protect individual's privacy* in this context are major challenges that need to be addressed in the future. There are users that do not expect privacy for any data they post on their social network; users that for minor financial benefits will voluntarily give up their private information; as well as users that are very privacy aware. To create a common view of privacy is a challenging task that

needs to be solved from a sociological perspective. Related challenges include *users' awareness of privacy issues* and *difficulty to create useful privacy legislation in an online medium* where users "voluntarily" provide sensitive information.

To that end, privacy in social networks requires a clear and near universal definition that can be updated through time. We need a standard model (perhaps similar to the Open System Interconnection (OSI) Reference model) for privacy in social networks. This model will address questions regarding the minimum acceptable requirements for a social network to be considered safe. This will be in terms of privacy dealing with each layer that contains or transports private information. It will require research into what today's social media consumers want as well as legislative aspects associated with privacy. Research should also address which predefined relationships for users of social networks bearing various privacy settings (e.g., just like those currently in existence on Facebook) should be encouraged to exist in social networking services by default. In addition, another important research issue is: should users have the power of customize relationships associations and their respective privacy settings and what is the degree of effectiveness for doing so? The literature indicates that people have a tendency to share private information even when they express privacy concerns [72]. Research needs to address whether user-based privacy customization is effective at protecting individual privacy.

There are also many technical challenges in social network privacy that provide opportunity for future research.

Balancing privacy and data utility remains an important challenge in this field. While much work has been done with respect to this problem, we still do not know how to share private data while protecting privacy and ensuring sufficient data utility in the shared data. The trade-off between utility and privacy was introduced in the form of the R–U confidentiality map [22]. Such a map is a set of values, R and U, of disclosure risk and data utility that correspond to various strategies for releasing the data. An example of such a map is shown in Fig. 12.11. Most of the work to release anonymized social networks is based on maximizing data utility while maintaining the disclosure risk under a given threshold. This technique, also known as privacy-based approach, corresponds to the RU map shown in Fig. 12.11.

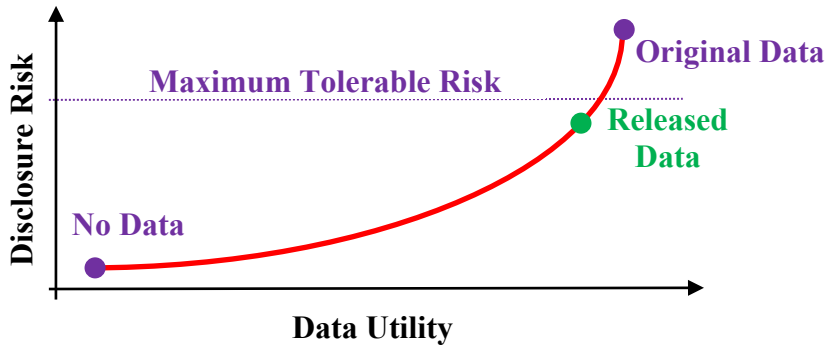


Fig. 12.11 An R–U confidentiality map

Social network anonymization still provides an imperfect solution. The availability of data from various sources makes anonymization more and more challenging. Finding more secure anonymization approaches while preserving data

utility remains a challenge in data privacy. The new paradigm of differential disclosure is promising but it requires better solutions that preserve network data utility to a satisfactory level. New solutions are needed for releasing data that are both confidential and preserve data utility.

Social networks are dynamic and protecting the individuals in this context is very challenging. Existing methods do not perform well with multiple releases of the data, because the data evolves in time and releasing just one version of the data is not acceptable in many practical problems. While there is some preliminary work in this area [78] more research is needed.

The advent of Big Data represents a privacy challenge as well. Businesses are able to use Big Data to learn more about their employees, increase productivity, and reduce cost. However, in these processes, the privacy of individuals is at high risk due to the high level of monitoring. Balancing how to use Big Data while preserving the privacy of individuals is a difficult problem that requires future research. Related to Big Data, the increasing use of technology generate more individual data. For instance, the use of wearable devices such as heart rate monitors or exercise devices and online activity (social networks, online searches, blogs) creates a continuous flux of data. To add to that, advances in Big Data analytics and other fields will likely reveal new trends and patterns about individuals. For instance, the likelihood of specific diseases such as Alzheimer may be computed in the near future based on genetic screening and other factors. Such developments will also create more privacy challenges.

The richness of information embedded in social networks creates major privacy challenges. A social network contains a variety of data in addition to its network structure. For instance geolocation data can be included as part of the profile, multimedia files may also contain sensitive information that is hard to detect without human intervention. How to protect individuals' privacy in this environment is extremely challenging and future research needs to address this problem.

An important opportunity that exists in this area is to the creation of privacy software tools. We envision two types of software tools that have the potential to increase the awareness of privacy issues and to make privacy more user friendly. In the first category of such tools, the social network users should automatically set their privacy preferences in a variety of social network sites. These tools have the potential to improve the social privacy component illustrated in Fig. 12.3. A second category of tools, used by social network owners, will aim to create anonymized social networks based on specified parameters. While prototypes of such tools exist for specific anonymization models, there are no tools that allow selection of the desired anonymity model and that are easy to use. Creating such privacy software tools will contribute to automating institutional privacy and in particular the network privacy component (see Fig. 12.3).

Finally, *privacy needs to be connected with deception literature and deception detection and prevention research*. Protecting one's privacy involves safeguarding software as much as safeguarding people from people. Social engineering has become prevalent through social networking sites [28], so privacy should not be examined disconnected from deception. Deception detection algorithms can contrib-

ute to helping maintain one's privacy by eliminating the potential for identity theft and consequences arising from that theft. Educating developers and designers as well as users about privacy also means educating them about deception. These two terms are linked. It is as necessary that we explore new research directions as that we update technical procedures that govern the development of social networking services.

References

[1] Alufaisan Y. and Campan A. 2013. Preservation of centrality measures in anonymized social networks.*Proceedings of the ASE/IEEE International Conference on Privacy, Security, Risk, and Trust (PASSAT 2013)*, Washington D.C., USA.

[2] Barbaro M. and Zeller T. 2006. A face is exposed for AOL searcher no. 4417749. *The New York Times*, Published August 9, 2006.

[3] Barnes S. B. 2006. A privacy paradox: social networking in the United States. *First Monday*,Vol.11, No.9, pp. 11–15.

[4] Beato F., Kohlweiss M., and Wouters K. 2011. Scramble! yoursocial network data. *Privacy Enhancing Technologies Symposium (PETS)*, pp. 211–225.

[5] Bernal J. 2009. Web 2.0 and social networking for the enterprise: guidelines and examples for implementation and management within your organization. *Pearson Education*.

[6] Bhat C. S. 2008. Cyber bullying: Overview and strategies for school counsellors, guidance officers, and all school personnel. *Australian Journal of Guidance and Counseling*, Vol. 18, No. 1, pp. 53–66.

[7] Blocki J., Blum A., Datta A., and Sheffet O. 2013. Differentially private data analysis of social networks via restricted sensitivity. *Proceedings of the Conference on Innovations in Theoretical Computer Science (ITCS)*, pp. 87–96. doi:10.1145/2422436.2422449

[8] Boyd D. M. 2003. Reflections on Friendster, trust and intimacy. *Proceedings of the Fifth International Conference on Ubiquitous Computing (Ubicomp 2003), Workshop application for the Intimate Ubiquitous Computing Workshop*,Seattle WA, USA.

[9] Boyd, D. M. 2004. Friendster and publicly articulated social networking. *Proceedings of the ACM CHI 2004 Conference on Human Factors in Computing Systems*, pp. 1279–1282. New York NY, USA. ACM Press. doi:10.1145/985921.986043

[10] Boyd D. M. and Ellison N. B. 2007. Social network sites: definition, history, scholarship. *Journal of Computer-Mediated Communication*, Vol. 13, No. 1, pp. 1–19.

[11] Brenner J. and Smith A. 2013. 72% of online adults are social networking site users. Available online at:http://pewinternet.org/Reports/2013/social-networking-sites.aspx

[12] Brown J., Broderick A. J., and Lee N. J. 2007. Word of mouth communication within online communities: conceptualizing the online social network. *Journal of Interactive Marketing*, Vol. 21, No. 3, pp. 2–20. doi:10.1016/10.1002/dir.20082

[13] Campan A. and TrutaT.M. 2008. A clustering approach for data and structural anonymity in social networks.*Proceedings of the 2nd ACM SIGKDD International Workshop on Privacy, Security, and Trust in KDD (PinKDD).*

[14] Chen R., Fung B., Yu P. S., and Desai B. 2014. Correlated network publication via differential privacy. *The VLDB Journal*, Vol. 23, Issue 4, pp. 653–676. doi:10.1007/s00778-013-0344-8

[15] Cheng J., Fu A. W. C., and Liu J. 2010. K-isomorphism: privacy preserving network publication against structural attacks. *Proceedings of the ACM SIGMOD International Conference on Magement of Data (SIGMOD)*, pp. 459–470.

[16] Choi D. and Kim J. 2004. Why people continue to play online games: in search of critical design factors to increase customer loyalty to online contents. *Cyberpsychology and Behavior*, Vol. 7, No. 1, pp. 11–24.

[17] Coleman J. S. 1988. Social capital in the creation of human capital. *American Journal of Sociology*, Vol. 94, pp. 95–120. doi:10.1086/228943

[18] Consumer Reports. 2012. Facebook and your privacy. Who sees the data you share on the biggest social network? *Consumer Reports Magazine*, June 2012.

[19] Dey R., Jelveh Z., and Ross K. W. 2012. Facebook users have become much more private: alarge-scale study.*PerCom Workshops*, pp. 346 – 352.

[20] Dingledine R., Mathewson N., and Syverson P. 2004. Tor: the second generation onion router. *USENIX Security Symposium*, pp. 303 – 320.

[21] DiNucci D. 1999. Fragmented future. *Print*, Vol. 53, No. 4, p. 32.

[22] Duncan G. T., Keller-McNulty S. A., and Stokes S. L. 2001. Disclosure risk vs. data utility: the R-U confidentiality map. *Technical Report Number 121*. National Institute of Statistical Sciences.

[23] Dwork C. 2006. Differential privacy. *Proceedings of the International Colloquium on Automata, Languages, and Programming (ICALP)*.

[24] Editorial Board. 2014. Edward Snowden, whistle-blower. *The New York Times*, Published January 2, 2014.

[25] Ellison N. B., Steinfield C., and Lampe C. 2007. The benefits of Facebook "friends:" social capital and college students' use of online social network sites. *Journal of Com-*

puter-Mediated Communication, Vol. 12, No. 4, pp. 1143–1168. doi:10.1111/j.1083-6101.2007.00367.x

[26] Facebook. 2014. Data use policy. Available online at: https://www.facebook.com/about/privacy/

[27] Goel V. 2013. Facebook to update privacy policy, but adjusting settings is no easier.*The New York Times*, Published August 29, 2012.

[28] Gross, R. andAcquisti, A. 2005. Information revelation and privacy in online social networks. *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pp. 71-80. doi:10.1145/1102199.1102214

[29] Guha S., Tang K., and Francis P. 2008. Noyb: privacy in online social networks. In *Proceedings of the First Workshop on Online Social Networks*, pp. 49 – 54.

[30] Guillen M. F. and Suarez S. L. 2005. Explaining the global digital divide: economic, political and sociological drivers of cross-national internet use. *Social Forces*, Vol. 84, No. 2, pp. 681–708. doi:10.1353/sof.2006.0015

[31] Gundry J. 1992. Understanding collaborative learning in networked organizations. *A. R. K. Ed (Ed.),Collaborative Learning through Computer Conferencing*, pp. 167–178. Berlin: Springer-Verlag.

[32] Gundry J. 2006. Web 0.0 social media. *Knowledge Ability Ltd*. Available online at: http://www.knowab.co.uk/socialmedia.html

[33] Gupta A., Roth A., and Ullman J. 2012. Iterative constructions and private data release. *Proceedings of the Theory of Cryptography Conference (TCC)*, pp. 339 – 356.

[34] Guynn J. 2012. Facebook changes privacy controls again and takes a key one away.*Los Angeles Times*, Published December 12, 2012.

[35] Hargittai E. 2008. Whose space? Differences among users and non-users of social network sites. *Journal of Computer-Mediated Communication*, Vol. 13, No. 1, pp. 276–297. doi:10.1111/j.1083-6101.2007.00396.x

[36] Hay M., Miklau G., Jensen D., Weis P., and Srivastava S. 2007. Anonymizing social networks. *Technical report, University of Massachusetts, Amherst.*

[37] Hay M., Li C, Miklau G., and Jensen D. 2009. Accurate estimation of the degree distribution of private networks. *Proceedings of the International Conference on Data Mining (ICDM).*

[38] Hoegg R., Martignoni R., Meckel M., and Stanoevska K. 2006. Overview of business models for Web 2.0 communities.*Proceedings of the GeNeMe (Gemeinschaften in NeuenMedien)*, Dresden, Germany, pp. 23–37.

[39] Howard T. W. (2010). Design to thrive: creating social networks and online communities that last. *Morgan Kaufmann*.

[40] Izquierdo E. 2011. Social networked media: advances and trends. *Proceedings of the 2011 ACM Workshop on Social and Behavioural Networked Media Access*, pp. 1–2. New York, NY, USA: ACM. doi:10.1145/2072627.2072629

[41] Jabeur N., Zeadally S., and Sayed B. 2013. Mobile social networking applications. *Communications of the ACM*, Vol. 56, No. 3, pp. 71–79. doi:10.1145/2428556.2428573

[42] Jones R. 1994. Digital's world-wide web server: a case study. *Computer Networks and ISDN Systems*, Vol.27, No. 2, pp. 297–306. doi:10.1016/0169-7552(94)90144-9

[43] Kaplan A. M., and Haenlein M. 2010. Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, Vol. 53, No. 1, pp. 59–68. doi:10.1016/j.bushor.2009.09.003

[44] Kaplan A. M. and Haenlein M. 2011. The early bird catches the news: nine things you should know about micro-blogging. *Business Horizons*, Vol. 54, No. 2, pp. 105–113. doi:10.1016/j.bushor.2010.09.004

[45] Karwa V., Raskhodnikova S., Smith A., and Yaroslavtsev G. 2011. Private analysis of graph structure. *Proceedings of the VLDB Endowment*, Vol. 4, No. 11, pp. 1146–1157.

[46] Karwa V. and Slavkovic A. 2012. Differentially private graphical degree sequences and synthetic graphs. *Proceedings of the Privacy on Statistical Databases Conference, Lecture Notes in Computer Science*, Vol.7556, pp. 273–285.

[47] Kasiviswanathan S., Nissim K., RaskhodnikovaS., and Smith A. 2013. Analyzing graphs with node differential privacy. *Proceedings of the Theory of Cryptography Conference (TCC)*, pp. 457–476.

[48] Katz J. E., Rice R. E., and Aspden P. 2001. The Internet, 1995-2000: access, civic involvement, and social interaction. *American Behavioral Scientist*, Vol. 45, No. 3, pp. 405–419. doi:10.1177/0002764201045003004

[49] Kietzmann, J. H., Hermkens, K., McCarthy, I. P., and Silvestre, B. S. 2011. Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, Vol. 54, No. 3, pp. 241–251. doi:10.1016/j.bushor.2011.01.005

[50] Kim A. J. 2000. Community building on the Web. *Peachpit Press*.

[51] Kirkpatrick D. 2010. The Facebook Effect. *Simon and Schuster*.

[52] Lane N., Walton-Flynn N., and Benlamlih F. 2008. Mobile social networking. *Informa UK Limited*. Available online at: http://www.telecoms.com/files/2009/05/buongiorno_final-fmt_nl-3110-f.pdf

[53] Lenhart A. and Madden M. 2007. Teens, privacy and online social networks: How teens manage their online identities and personal information in the age of MySpace.

Pew Internet and American Life Project. Available online at: http://apo.org.au/?q=node/16750

[54] Liu K. and TerziE. 2008. Towards identity anonymization on graphs.*Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data*, pp. 93–106.

[55] Lucas M. andBorisov N. 2008. Flybynight: mitigating the privacy risks of social networking. *Proceedings of the 7th ACM Workshop on Privacy in the Electronic Society (WPES)*, pp. 1–8.

[56] Luo W., Xie Q., and Hengartner U. 2009. FaceCloack: an architecture for user privacy on social networking sites. *Proceedings of the International Conference on Computational Science and Engineering (CSE)*, Vol. 3, pp. 26–33.

[57] Morris M. G. andVenkatesh V. 2000. Age differences in technology adoption decisions: implications for a changing work force. *Personnel Psychology*, Vol. 53, No. 2, pp. 375–403. doi:10.1111/j.1744-6570.2000.tb00206.x

[58] Muramatsu J. and Ackerman M. 1998. Computing, social activity, and entertainment: afield study of a game MUD. *Computer Supported Cooperative Work (CSCW)*, Vol. 7, No. 1-2, pp. 87–122. doi:10.1023/A:1008636204963

[59] Murchu I. O., Breslin J. G., and Decker S. (2004). Online social and business networking communities. *Proceedings of the ECAI 2004 Workshop on Application of Semantic Web Technologies to Web Communities*, pp. 241–267. doi:10.1007/978-1-4419-7142-5

[60] Murugesan S. 2007. Understanding Web 2.0.*IT Professional*, Vol. 9, No. 4, pp. 34–41. doi:10.1109/MITP.2007.78

[61] Nobari S., Karras P., Pang H., and Bressan S. 2014. L-Opacity: Linkage-Aware Graph Anonymization. *Proceedings of the International Conference on Extending Database Technology (EDBT)*, pp. 583–594.

[62] O'Brien C. N. 2011. The first Facebook firing case under Section 7 of the National Labor Relations Act: exploring the limits of labor law protection for concerted communication on social media. *Suffolk University Law Review*, Vol. 45, pp. 29–66.

[63] Olson P. 2013. Teenagers say goodbye to Facebook and hello to messenger apps. *The Observer Journal*, Saturday 9 November 2013, Online at: http://www.theguardian.com/technology/2013/nov/10/teenagers-messenger-apps-facebook-exodus

[64] QuerciaD., Lathia N., Calabrese F., Di Lorenzo G., and Crowcroft J. 2010. Recommending social events from mobile phone location data. *Proceedings of the IEEE 10th International Conference on In Data Mining (ICDM)*, pp. 971–976. doi:10.1109/ICDM.2010.152

[65] Rafaeli S. and Larose R. J. 1993. Electronic Bulletin Boards and "Public Goods" Explanations of Collaborative Mass Media. *Communication Research*, Vol. 20, No. 2, pp. 277–297. doi:10.1177/009365093020002005

[66] Raynes-Goldie K. S. 2012. Privacy in the age of Facebook: discourse, architecture, consequences. *Ph. D. Thesis*, Curtin University, January 2012.

[67] Rome Memorandum. 2008. Report and Guidance on Privacy in Social Networks Services – "Rome Memorandum", *International Working Group on Data Protection in Telecommunications*, Rome, Italy.

[68] Rosenblum D. 2007. What anyone can know: the privacy risks of social networking sites.*IEEE Security Privacy*, Vol. 5, No. 3, pp. 40 – 49. doi:10.1109/MSP.2007.75

[69] Shen X. (Sherman). 2013. Security and privacy in mobile social network [Editor's Note]. *IEEE Network*, Vol. 27, No. 5, pp. 2–3. doi:10.1109/MNET.2013.6616107

[70] Shuen A. 2008. Web 2.0: astrategy guide. *O'Reilly Media, Inc*.

[71] Sicilia M. andPalazón M. 2008. Brand communities on the internet: a case study of Coca-Cola's Spanish virtual community. *Corporate Communications International Journal*, Vol. 13, No. 3, pp. 255–270. doi:10.1108/13563280810893643

[72] Squicciarini, A. C. and Griffin, C. 2012. An informed model of personal information release in social networking sites. *Proceedings ofthe 2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*, pp. 636–645. doi:10.1109/SocialCom-PASSAT.2012.137

[73] Squicciarini A., Shehab M., and Wede J. 2010. Privacy policies for shared content in social network sites. *The VLDB Journal*.

[74] Stutzman F., Gross R., and Acquisti A. 2012. Silent listeners: the evolution of privacy and disclosure on Facebook. *Journal of Privacy and Confidentiality*, Vol. 4, No. 2, pp. 7–41.

[75] Tassa T. and Cohen D. J. 2013. Anonymization of centralized and distributed social networks by sequential clustering. *IEE Transactions on Data and Knowledge Engineering*, Vol. 25, Issue 2, pp. 311–324. doi:10.1109/TKDE.2011.232

[76] Travers J. and Milgram S. 1969. An experimental study of the small world problem. *Sociometry*, Vol. 32, No. 4, pp. 425–443. doi:10.2307/2786545

[77] Valkenburg P. M., Peter J. and Schouten A. P. 2006. Friend networking sites and their relationship to adolescents' well-being and social self-esteem. *CyberPsychology Behavior*, Vol. 9, No. 5, pp. 584–590. doi:10.1089/cpb.2006.9.584

[78] Wang C.-J. L., Wang E. T., and Chen A. L. P. 2013. Anonymization for multiple released social network graphs. *Advances in Knowledge Discovery and Data Mining*, LNCS Volume 7819, pp. 99–110. doi: 10.1007/978-3-642-37456-2_9

[79] West A., Lewis J., and Currie P. 2009. Students' Facebook "friends": public and private spheres. *Journal of Youth Studies*, Vol. 12, No. 6, pp. 615–627. doi:10.1080/13676260902960752

[80] Yamada A., Kim T. H., and Perrig A. 2012. Exploiting privacy policy conflicts in online social networks. *CMU-CyLab-12-005*, Carnegie Mellon University.

[81] Zheleva E. and Getoor L. 2011. Privacy in social networks: a survey. Chapter in Social Network Data Analytics, *Springer Science and Business Media*.

[82] Zhou B. and PeiJ. 2008. Preserving privacy in social networks against neighborhood attacks.*Proceedings of the IEEE International Conference on Data Engineering (ICDE)*, 506-515.

[83] Zou L., Chen L., and Ozsu T. M. 2009. K-automorphism: a general framework for privacy preserving network publication. *Proceedings of the International Conference on Very Large Data Bases (VLDB)*.