Privacy Implications of Privacy Settings and Tagging in Facebook*

Stan Damen and Nicola Zannone

Eindhoven University of Technology
s.damen@student.tue.nl
n.zannone@tue.nl

Abstract. Social networks are becoming increasingly popular nowadays. Users share personal information about themselves and other users in order to build and maintain their social network. However, the large amount of personal information available on social networks poses risks of data misuse. Although social networks offer users the possibility to specify privacy settings to regulate access to their information, these settings are often complicated and unintuitive, especially when dealing with new modalities of social communication like tagging. In this paper we investigate the privacy consequences of information sharing in social networks. In particular, we formally analyze the impact of the privacy settings and the use of tagging in Facebook on the visibility of information. To increase users' awareness of the risks of information sharing and empower users to control their information, we present a tool for determining the visibility of users' information based on their privacy settings and tagging.

1 Introduction

Online social network services, also called social networks, have become increasingly popular over the years. For instance, social networks like Facebook, Google+ and Twitter have millions of users across the world. The popularity of social networks is due to the fact that people want to keep in contact with their friends and meet people with common interests. Social networks provide a social environment in which users can share information with other users and build communities around common interests.

The most common way to share information is in the form of posts which can be placed by users on their own profile or on the profile of other users. Other examples include the possibility of sharing pictures, having profiles that are (partially) publicly available, and options to provide additional information about the user (e.g., location information). Many social networks also allow users to add to their profile (third party) applications which provide additional functionalities (e.g., games, online marketplaces, function enhancers) for sharing information and building their social network.

Social networks have also led to the introduction of new modalities of social communication for sharing information and building online communities. A prominent example of such new modalities is tagging which has been introduced by Facebook in 2009 [17]. Tagging allows users to share contextual information about themselves or

* This work is funded by the Dutch national program COMMIT through the THeCS project.

their friends by linking a user to a certain content on the social network. In particular, a tag is a label specifying a user's name and provides a link to that user's profile.

From a user's viewpoint, the uncontrolled sharing of personal information poses potential privacy threats [18,23]. In particular, information available on social networks can be misused by other users (e.g., cyberstalking [11], identity theft [4,27] and discrimination [12]). In order to address user privacy concerns, many social networks allow users to specify privacy settings in order to regulate the visibility of their information. In addition, tools are available within social networks to help users visualize their social circle and the visibility of the information posted in their profile.

Although privacy settings provide users some control over their information, such settings are often complicated and unintuitive. In particular, they may mislead users by providing confidence to be in full control of their information. Many users believe they are solely sharing data with their friends and are unaware that the actual visibility may not reflect their privacy settings [32]. For instance, tags modify the visibility of objects, making it difficult for users to determine to what extent a piece of information is visible. Moreover, tools for viewing the user's profile from the perspective of other users often do not reflect the real visibility of information. As a consequence, they provide a false perception that leads users to underestimate the risks of sharing information.

Another main privacy issue concerns the user who is in control of the information. In social networks, the user in control of the information is usually the user who owns the profile in which the content is posted. In contrast, privacy regulations (e.g., Directive 95/46/EC and its subsequent regulation) empower the data subject – i.e., the user to whom the information refers – to control the processing and disclosure of his data [9].

To enable users to control the use of their data, we need collaborative access control systems able to support the functionalities provided by social networks. Moreover, these systems should increase users' awareness of the privacy risks of sharing information. In particular, they should assist users in ensuring that the specified settings reflect their intentions and in understanding the privacy consequences of sharing information.

This work takes a first step in the development of such systems. We formally analyze the impact of privacy settings and tagging on the visibility of information in Facebook and identify drawbacks in the privacy controls used to regulate access to information. First, we model user profiles in Facebook along with the objects that can be shared by users as well as the role of users with respect to information. We use the profile model to study how the visibility is determined by privacy settings and tagging. In particular, the model has been used to develop a proof-of-concept tool which aims to increase awareness and empower users to control their information. The tool implements Facebook's privacy controls in Prolog, and allows users to determine the visibility of their information based on their settings and tags. To make the discussion more concrete, we analyze a number of scenarios that are representative for real situations in Facebook.

The paper is organized as follows. Section 2 discusses the privacy issues in social networks. Section 3 presents the Facebook profile model, and Section 4 demonstrates the effect of privacy settings and tagging on the visibility of information using some examples. Section 5 presents an implementation of Facebook's privacy controls in Prolog to determine and analyze the visibility of information. Finally, Section 6 discusses related work, and Section 7 concludes the paper providing directions for future work.

2 Privacy Issues in Social Networks

To build and maintain their social circle, users of social networks are willing to share more and more information about themselves and about other users. Larcom and Elbirt [13] observe that *"the important common thread among these* [social network] *services is the exchange of personal information over the Internet"*. Thus, a huge amount of personal data is available on social networks nowadays. Although the sharing of personal data helps users build large social circles, this attitude poses privacy risks to them.

Several studies [10,15,18,23,25] have analyzed privacy concerns in social networks. Privacy issues in social networks can be classified into two categories.

Social network privacy practices: this category concerns privacy issues related to the collection and processing of personal data by the social network and their disclosure to third parties. Privacy issues in this category include user tracking (e.g., Facebook "Like" button [16]), user profiling for advertisement purposes and secondary usage of data [22], and storing information after it was deleted by the user.

Information disclosure to contacts: this category concerns privacy issues that arise from the misuse of personal information by other users in the social network. Privacy issues like cyberstalking [11], identity theft [4,27] and discrimination [12] fall under this category.

The first category is similar to the issues characterizing other domains in which personal data are handled by an organization. In this paper, we focus on the second category which is specific to social networks and, in general, to collaborative environments.

Users usually share their personal information on social networks voluntarily. Atwan and Lushing [2] observe that: *"There is only one thing in the world worse than being Facebook stalked, and that is not being Facebook stalked"*. This privacy paradox shows the contradictory desires of users: on the one hand, users want their privacy protected; on the other hand, they are willing to share more and more information about themselves in order to build and maintain their social network. Most users sacrifice their privacy in favor of their sociability. This choice is due to the fact that users are often not fully aware of the risks of sharing their information and of the widespread accessibility of information when posting on social networks.

Social networks allow users to specify privacy settings to control the visibility of the objects in their profile (i.e., who can see the object). This, however, may mislead users into believing they are in full control of their information. Even if users specify their privacy settings carefully, the information can be viewed by more users than the profile owner intended. For instance, in Facebook tagging a post modifies the visibility of the post. Moreover, by tagging a post or an image, a copy of the tagged post or image appears in the profile of the tagged user. The latter can then specify the visibility of the copy in his profile, regardless of the privacy settings of the original post.

The first intuition for analyzing the privacy issues in social networks and, in particular, the control over the information is that we need to distinguish the roles of users involved with the management of information. The profile owner has control of the information posted on his profile and in particular on its visibility. When posting on the profile of another user, a user retains some rights over the posted information. Tagged users have control over the tag and can influence the visibility of the information to

which they are tagged. Last but not least, users can share information not only about themselves, but also about other users. Data subjects should be able to control the information about them and in particular its visibility. In the remainder of the paper, we refer to the problem of controlling the usage of information as the *multi-ownership* problem.

Multi-ownership introduces a number of privacy risks as it is not easy to understand to what extent users can control information, especially when users involved in the management of a piece of information specify conflicting privacy settings or are not aware of the privacy settings specified by the data host and of his relationship with other users. For instance, a user can establish a friend relationship with another user just to become close to a third user and therefore access the information in his profile without the latter knowing it. In addition, users may not be aware of the existence of content concerning themselves in the profiles of other users. The main issue here is that it is very difficult, if not impossible, to correctly identify the users to whom the information refers. Even if users would be aware, they have very little control over their information posted in the profile of other users. In particular, they have no authority to remove their information from other profiles or share it with a smaller group of users.

In this paper, we analyze the problem of multi-ownership and the issues introduced by the tagging functionality in Facebook. In particular, by formalizing the privacy settings in Facebook, we aim to study how tagging affects the multi-ownership problem and influences the visibility of information.

3 Facebook Profile Model

Facebook allows users to define privacy settings to control the usage of the information published in their profile (visibility, posting, removal, etc.). In the remainder of the section, we describe the information contained in a Facebook profile and the permission that users have on such information. The Facebook profile model is presented in Fig. 1.

The main objects on Facebook are user profiles. Users can be individuals or organizations. A *profile* is created when a user signs up for the social network, and can be seen as a representation of the user. In particular, in Facebook a user is linked to a single profile and a profile to a single user. Therefore, we freely interchange the terms user and profile from here on. Every profile has a standard set of *profile information* associated with it (e.g., name, country, email address), and can contain one or more albums and posts. An *album* is a collection of *images*. A *post* is a message published in a profile. The main difference between images and posts is that images can only be uploaded on a user's own profile, while posts can also be placed on the profile of other users. Note that a user can post an image on the profile of another user; however, in this case the image is considered a post. Posts and images can have one or more comments and tags connected to it.[1] A *comment* is a note commenting a target object. A *tag* is a label that links a piece of information to a user profile. The concept of "wall", which is used on Facebook as a central place where content is visible, is not modeled since privacy settings cannot be defined at wall level.

[1] Although Facebook also allows users to include tags in comments, we do not consider such tags in the paper as they do not change the visibility of objects.

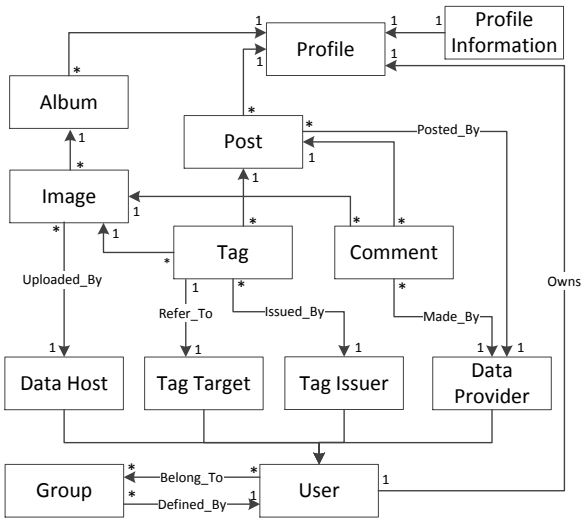


Fig. 1: Facebook Profile Model

We distinguish four roles that users can have in Facebook. The *data host* is the user that owns the profile in which the content is posted. The *data provider* is the user placing content on a profile in the form of posts, comments and images. The *tag issuer* is the user placing a tag on a post or an image. The *tag target* is the tagged user. Note that users can have more than one role on an object, e.g. data host and data provider.

To support users in organizing and managing their online relations, social networks employ the concept of *group*. A group consists of a set of users. Facebook provides users with a number of standard groups (i.e., friends, friends of friends, public, only me) as well as with the possibility to create custom groups specific to users' needs (e.g., school friends). Note that groups are specific to a profile, i.e., the membership of a user to a group should be determined with respect to the user who has defined the group. This is represented by cardinality "1 to many" for relation *defined_by* in Fig. 1.

The permissions that users can have on an object depend on their role on the object. The permissions supported by Facebook are "view", "delete", "post", "comment" and "tag". Note that permission to "comment" is implied by permission to "view"; permission to "tag" is implied by permission to "post" for posts and by permission to "view" for images. Therefore, we will not consider these permissions further in the paper.

The data host has all permissions on his profile and on the objects in it. The data provider can remove the content he posted at the condition he is in the current visibility of the object: if the data host changes his setting and the object is no longer visible to the data provider, then the data provider cannot remove the object anymore. Finally, the tag target has permissions to remove the tag from the content, regardless of its location.

Facebook also supports the "control" permission, which represents the authorization to change profile settings. Such permission is always limited to the data host.

The permissions a user has on an object are also based on the visibility of objects. The visibility of an object is determined by its location, the privacy settings defined for the object and the tags associated to it. In Facebook, privacy settings can be specified at different levels of granularity, namely profile, profile information, album, post and image. The visibility of an object consists of the group defined by the data host in his settings for the object. Intuitively, a user can view the content in the profile of another user only if he belongs to the group to which the profile owner has granted permission to view the content. In addition, for every tag attached to a post, the same type of group, now defined by the tag target, is added to the visibility of the post. When a post or an image is tagged, a "copy" of the object appears on the profile of the tagged user. The visibility of the copy is determined regardless of the settings for the original object.

Facebook also allows users to specify profile settings. In this paper, we only consider the *post setting*, which specifies which group can post on the profile. The options for this setting are limited to friends or "no one" (i.e., only the profile owner can post). The other settings are used to specify default groups for the visibility of new objects. In addition, a user can specify whether the objects in which he is tagged should appear on his profile. In the remainder of the paper, we assume that this option is active with "friends" as visibility (note that these are the default settings in Facebook). This makes it possible to study information dissemination, of which the data host is unaware.

4 Application of Privacy Settings

In this section, we analyze the multi-ownership problem and the effect of tagging through a number of scenarios that are representative for real situations in Facebook.

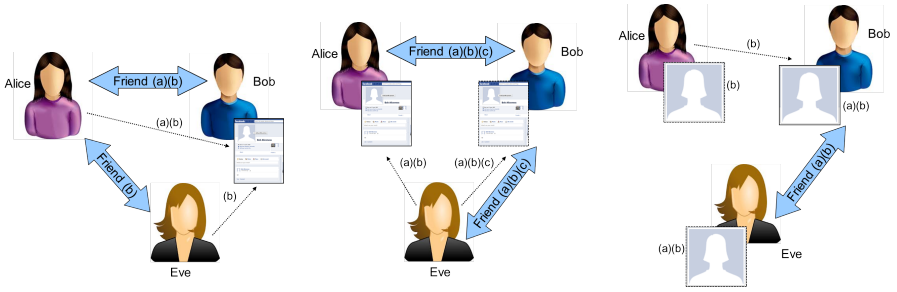
4.1 Scenario 1

This scenario analyzes a situation in which tagging is not used. It includes three users: Alice, Bob and Eve. Alice and Bob are friends. The scenario is shown in Fig. 2a.

Scenario 1(a) Alice posts some content on Bob's profile. Accordingly, Bob is the data host; Alice is the data provider. Bob assigns visibility *friends of friends* to the post. He can delete the post; Alice can delete the post as long as she is in the visibility of the post.

Scenario 1(b) Eve wants to view the information posted on Bob's profile without the latter knowing it. To this end, Eve becomes a friend of Alice. As the post has visibility *friends of friends*, Eve can view the post and all comments in response to it.

This simple scenario shows that, when using the standard posting feature of Facebook to share information, the data host is in full control of the content in his profile. One may argue that Bob may not know that Eve can view the post in his profile. However, Bob can simply avoid it by specifying more restrictive privacy settings. For instance, he can change the visibility of the post to *friends* or *only me*.



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3

Fig. 2: Evaluation Scenarios

4.2 Scenario 2

This scenario extends scenario 1 by considering the use of tagging. Here, Bob is friend of Alice and Eve. The scenario is shown in Fig. 2b.

Scenario 2(a) Alice posts some content about Bob on her profile and tags Bob. Accordingly, Alice is the data host and data provider; Bob is the tag target. The visibility of the post is set by Alice to *friends of friends*. Because of the tag, the visibility of the post also includes the *friends of friends* of Bob. In addition, a copy of the post appears on Bob's profile (Bob is the data host for the copy). The visibility of this copy is *friends* (of Bob) based on the default profile setting of Bob for tags. Because of the friendship relations between the users, Eve is part of the visibility of the post.

Scenario 2(b) Alice realizes Eve is in the visibility of the post and wants to stop sharing it with Eve while allowing her friends to view it. To do this, Alice changes the visibility to *friends*. As a consequence, the visibility of the post changes to Alice's *friends* and Bob's *friends*. Nonetheless, Eve remains part of the visibility, as she is Bob's friend.

Scenario 2(c) Since changing the visibility to *friends* did not work, Alice changes the visibility to *only me* to prevent Eve to view the post. Thus, the visibility of the post changes to only Alice and Bob. However, Eve can still see the content in the post, as she can see the copy on Bob's profile.

This scenario illustrates the privacy risks caused by the extended, uncontrollable visibility induced by tagging. Alice changed the visibility in an attempt to remove a specific user from the visibility. However, due to the complexity introduced by the tag, even changing the visibility to *only me* does not have the desired result. The problem is that Alice depends on Bob for the visibility of the post. At this point, the only option left to Alice would be to remove the tag. However, by removing the tag Bob loses any form of control on the post, which is unacceptable from a privacy perspective as he is the subject of the content and therefore he should retain some authority on it. The alternative of leaving the tag attached to the post is also unacceptable, as Alice cannot

restrict the visibility as desired. This simple example shows that the use of tagging in Facebook makes the multi-ownership problem, and control of the visibility, non-trivial.

4.3 Scenario 3

This scenario is based on scenarios 1 and 2, and aims to illustrate additional problems related to the use of tagging. Here, Bob and Eve are friends; Alice is not a friend of Bob and Eve. The scenario is shown in Fig. 2c.

Scenario 3(a) Bob uploads an embarrassing photo of Alice in an album that he shares with his friends. Bob tags Eve in the image, making an instance of the image appear on Eve's profile. The visibility of the image is Bob and Eve's *friends* (for the image in Bob's profile) and Eve's *friends* (for the image in Eve's profile). At this point, Alice may not be aware that a photo of her has been uploaded.

Scenario 3(b) A friend of Bob tags Alice to make sure everyone knows which person is the subject of the image.[2] Alice is notified that she has been tagged and, thus, she becomes aware of the existence of the image. In response, Alice deletes the image from her profile. However, the image still appears on Bob's profile, including the tag pointing to Alice. Alice decides to remove the tag as well. However, the image in Bob's profile and its copy in Eve's profile remain visible.

This scenario illustrates the lack of control users have over their information when posted on other users' profiles. Initially, Alice is not even aware that an image about her is posted (scenario 3(a)). She becomes aware in scenario 3(b) after she is tagged. However, she does not have the right to delete the image from other users' profile or even to restrict its visibility; she is only able to remove the image from her profile and the tag from the image.

4.4 Discussion

To analyze the consequence of information sharing, users require a good understanding of how the visibility of information is determined. However, privacy controls in Facebook are complicated and unintuitive. In particular, the scenarios above show that determining the visibility of objects becomes increasingly complex when tagging is used. For instance, tagging modifies the visibility of posts by including the tagged user's group corresponding to the one specified by the data host in his settings. As a consequence, users may share their information with more users than they intended.

Moreover, tags create a copy of tagged objects in the profile of the tag target. Facebook adopts an object-centric approach in which copies are treated as individual objects: users can define the visibility of copies in their profile regardless of the privacy settings for the original objects. This makes it difficult for the data host of the original object to restrict the visibility to certain users. For instance, scenario 2 shows that the only option Alice has to completely remove Eve from the visibility is to remove the tag.

[2] Bob's friend and Alice have to be friends as this is a requirement for tagging in Facebook.

To determine the actual visibility of the content (as opposed to objects) the data host of the original object needs to know the settings defined by the tagged user as well as his relations with other users. This, however, is impossible in Facebook as no one but the data host can visualize his settings. Although Facebook's choice of keeping settings private is reasonable, the settings for the original object should be considered when calculating the visibility of its copies.

The identification of data subjects and the control they have over their information are other crucial issues in social networks. Although tagging may be seen as a solution to the problem of linking a piece of information to the corresponding data subject, tagged users are not necessarily related to the information (see scenario 3). Indeed, the main goal of tagging in Facebook is to make information easy to access rather than identifying the actual data subject(s). As a result, tagging can grant some control over the information to users that are not directly related to the information, increasing the risks of data misuse.

Even if data subjects are correctly tagged, they have limited control over their information (see scenario 3(b)). This is because Facebook assumes that the data host is the owner of the information. However, this is not always the desired solution. For instance, in scenario 3, it should be Alice in control of the visibility of the image; for Bob, delete permission would be sufficient. The obvious difficulty lies in determining the correct permissions for each piece of information, which is a non-trivial problem. However, making the data host automatically the owner is not a viable solution in all cases.

Another issue related to tagging is that when a user is tagged, the object automatically appears in her profile. The user may prevent it to occur by modifying the profile settings; however, this is a "one-size fit all" solution: either all contents in which she is tagged appear on her profile or none. A more desirable solution would be to let the tagged user pre-approve content before it appears on her profile.

5 Visibility Visualization Tool

In this section, we present an implementation of Facebook's privacy controls in Prolog for determining the visibility of users' information based on their privacy settings and tagging (available at http://security1.win.tue.nl/THeCS/). This proof-of-concept tool aims to increase awareness of the risks of information sharing and empower users to control their information.

5.1 Formal Representation of Privacy Settings

We use Prolog [1] to model and reason on the visibility of information based on user settings and tagging. First, we recap the Prolog concepts that are relevant to this paper.

An *atom* is an object of the form $p(t_1, \ldots, t_n)$ where p is an n-ary predicate symbol and $t_1, \ldots, t_n$ are terms (i.e., variables and constants). An atom is *ground* if $t_1, \ldots, t_n$ are constants. A *rule* is a construct of the form H :- $B_1, \ldots, B_n$ (with $n \geq 0$), where H is an atom called *head* and $B_1, \ldots, B_n$ (called *body*) are atoms. Intuitively, H is true if $B_1, \ldots, B_n$ are true. A *fact* is a rule with empty body (i.e., $n = 0$). A *program* is a finite set of rules.

Objects:
profile($profileID$)
profile-info($attributeID, profileID$)
album($albumID, profileID$)
image($imageID, albumID, tag\text{-}list$)
post($postID, profileID, poster, tag\text{-}list$)
comment($commentID, locationID, commenter$)
tag($tagID, issuer, target$)
Visibility:
setting($objectID, group$)
visibility-tag($objectID, visibility\text{-}list, tag\text{-}list$)
visibility($objectID, visibility\text{-}list$)
profile-setting($profileID, group$)
Membership:
friends($profileID_1, profileID_2$)
friendsOfFriends($profileID_1, profileID_2$)
belongsTo($profileID_1, (profileID_2, group)$)
member($profileID, visibility\text{-}list$)
Authorization:
can($profileID, permission, objectID$)

Table 1: Predicates

Table 1 shows the predicates used to represent settings and objects in the Facebook profile and to reason about them. Predicates profile, profile-info, album, image, post, comment and tag are used respectively to identify profiles, profile information, albums, images, posts, comments, and tags. The first argument of these predicates is the ID of the object and is used to identify the object itself. In addition, objects are linked to a higher level object; the highest level object is the profile. Profile information, albums and posts are linked to a profile, images to an album, and comments to a post or an image. Intuitively, this link is used to identify the user/profile hosting the object. Predicates post, comment and tag specify the data provider (e.g., poster, commenter) or tag issuer. Predicates post and image specify the list of tags associated to them. Finally, tag specifies the user that has been tagged.

To determine the visibility of an object, we employ three predicates: setting for representing the privacy settings defined by the data host for the object, visibility-tag for tag induced visibility, and visibility for representing the object's visibility based on settings and tags. Predicate setting is a binary predicate where the first argument is an object ID and the second is the group specified by the data host as privacy setting (e.g., friends, public). Predicate visibility is a binary predicate where the first argument is an object ID and the second is a visibility list. A visibility list is a list of pairs $(profileID, group)$; each pair specifies a group which is part of the object's visibility together with the user who defined the group. For example, the list $[(profile_1, friends), (profile_2, fof)]$ means that the friends of $profile_1$ and friends of friends of $profile_2$ form the object's visibility. The ternary predicate visibility-tag is similar to visibility; besides specifying the object ID and the visibility list of the object, it also provides the list of tags to be consid-

	Visibility
1	visibility(ID, X) :- comment$(ID, LocationID, _)$, visibility$(LocationID, X)$.
2	visibility$(ID, [(ProfileID, X)\|Y])$:- post$(ID, ProfileID, _, T)$, setting(ID, X), visibility-tag(ID, Y, T).
3	visibility-tag$(_, [], [])$.
4	visibility-tag$(ID, [(X, Z)\|Y], [TagID\|T])$:- visibility-tag(ID, Y, T), tag$(TagID, _, X)$, setting(ID, Z).
	Membership
5	friends(X, X) :- profile(X).
6	friends(X, Y) :- friends(Y, X).
7	friendsOfFriends(X, Z) :- friends(X, Y), friends(Y, Z).
8	belongsTo$(X, (Y, friends))$:- friends(X, Y).
9	belongsTo$(X, (Y, fof))$:- friendsOfFriends(X, Y).
10	belongsTo$(X, (Y, public))$:- profile(X), profile(Y).
11	member$(S, [(X, Y)\|Z])$:- belongsTo$(S, (X, Y))$.
12	member$(S, [T\|Z])$:- member(S, Z).
	Authorization
13	can$(S, delete, ID)$:- post$(ID, S, _, _)$.
14	can$(S, delete, ID)$:- comment$(ID, LocationID, _)$, post$(LocationID, S, _, _)$.
15	can$(S, delete, ID)$:- tag$(ID, LocationID, _)$, post$(LocationID, S, _, T)$, in(ID, T).
16	can$(S, delete, ID)$:- post$(ID, _, S, _)$, visibility(ID, X), member(S, X).
17	can$(S, delete, ID)$:- comment$(ID, _, S)$, visibility(ID, X), member(S, X).
18	can$(S, delete, ID)$:- tag$(ID, _, S)$.
19	can$(S, view, ID)$:- visibility(ID, X), member(S, X).
20	can$(S, post, ID)$:- profile-setting(ID, X), member(S, X).
	Copy Inference
21	copy$(ProfileID, PostID)$:- post$(PostID, _, _, Tags)$, tag$(TagID, _, ProfileID)$, in$(TagID, Tags)$.

Table 2: Rules for visibility, membership, and authorization

ered for determining the visibility of the object. In addition, we use the binary predicate profile-setting to specify post settings. The first argument is a profile ID, and the second denotes the group of users that can post on the profile.

The membership of a user to a group is determined using four predicates. Predicate member is used to determine whether a user is a member of the visibility list associated to an object. Predicate belongsTo is used to determine whether the user corresponding to $profileID_1$ is part of $group$ as defined by the user corresponding to $profileID_2$ (Remark that groups are defined with respect to users). Predicate friends$(profileID_1, profileID_2)$ holds if the user corresponding to $profileID_1$ is a friend of the user corresponding to $profileID_2$. Similarly, friendsOfFriends$(profileID_1, profileID_2)$ holds if the user corresponding to $profileID_1$ is a friend of a friend of the user corresponding to $profileID_2$. Finally, predicate can is used to determine whether a user has a given permission (i.e., *view*, *post* and *delete*) on an object. Intuitively, can$(profileID, permission, objectID)$ holds if the user corresponding to $profileID$ can exercise $permission$ on the object corresponding to $objectID$.

Table 2 provides the set of rules used to determine which permissions users have on a certain object. According to Prolog convention, we use symbol underscore (_) to denote an anonymous variable; intuitively, it means "any term". Rules 1 to 4 determine the visibility of posts and comments.[3] The visibility of comments depends on the visibility of the post to which a comment belongs. Accordingly, rule 1 associates to a comment the visibility list of the post to which the comment belongs. Rule 2 determines the visibility list of a post based on its privacy settings and the list of tags associated to it. The visibility list of the post implied by the tags associated to it is recursively built using rules 3 and 4. In particular, rule 4 sets the visibility of the post to the same group specified by the data host in her setting, but now defined by the tag target.

Rules 5 to 12 are used to determine the membership of a user to a group. In Facebook the friendship relation is both reflexive and symmetric. These properties are represented by rules 5 and 6, respectively. Rule 7 uses predicate friends to determine the friends of friends of a user. Rules 8, 9, and 10 determine the membership of a user to groups *friends*, friends of friends (*fof*), and *public*. The membership of a user to group *only me* as well as to custom groups can be explicitly specified using predicate belongsTo. Rule 8 states that a user belongs to the group *friends* of a certain profile if he is a friend of the user of that profile. Rule 9 is analogous to rule 8 but for group *fof*. Rule 10 states that, if a profile exists, then all existing profiles are part of group *public*. Rules 11 and 12 recursively verify whether a user is in the visibility list of an object.

Rules 13 to 20 are used to determine the permissions that a user has on an object. A user has *deletion* rights over a certain object in three cases. Rules 13 to 15 state that the data host has the right to delete the objects in his profile (e.g, posts, comments, and tags). Predicate in is used to determine if a tag is one of the tags contained in the post. Rules 16 and 17 state that a user can delete the posts and comments he gave only if he is still within the visibility of the post and comment, respectively. Finally, a user can delete the tags that point to him (rule 18). Note that the tag issuer cannot remove the tags he created. He can only delete the tag by deleting the post that contains it.[4] Finally, a user has the permission to view an object if he is in the visibility of the object (rule 19). The permission to post is determined by checking the profile setting for posting (rule 20).

To determine the visibility of a certain content it is not sufficient to determine the visibility of the object in which the content is posted; we also need to determine the visibility of the copies of the object due to tagging. This can be done by adding rules that, given a set of tags, infer the copies of the tagged object. Rule 21 shows how the copies can be inferred using the original post and the tags associated to it. Then, the visibility of the content contained in a post can be determined as the union of the visibility of the post and the visibility of its copies. This choice reflects the fact that Facebook treats copies as separated objects on which the corresponding data host specifies his settings. Note that deleting a copy can be done in two ways: either the user hosting the copy (i.e., the tag target) deletes it from her profile, or the data host or tag target removes the tag.

[3] The rules for images and albums are similar to the ones for posts.

[4] A tag can be inserted in a post only when the post is created. Accordingly, the tag issuer and data provider coincide for posts.

5.2 Proof-of-Concept

Facebook provides a functionality called "view as", which enables a user to look at her profile from the perspective of another user. In particular, using this functionality a user can determine the visibility of data objects in her profile from the perspective of a friend or a "public" user. However, this functionality only provides a partial view of the visibility. For instance, in scenario 1, Bob cannot verify whether Eve is in the visibility of the post. He can only verify the visibility that a public user has (Eve is not a friend of Bob), which shows that public users cannot view the information although Eve can. Moreover, the functionality does not consider tags to determine the overall visibility. In particular, the extended visibility introduced by adding tags to objects, as in scenario 2, is not shown; also copies of data objects are not considered for the visibility.

In contrast, the formalization of the Facebook profile model and privacy controls in Section 5.1 provides a visualization of the visibility which reflects the actual Facebook privacy controls. In particular, the proof-of-concept can accurately determine the visibility in situations involving tags and copies, like the scenarios in Section 4. A user can gain a more realistic insight into the access that others have to her information and therefore into the risks of information sharing by determining the visibility of her information, including instances of the information in other profiles. Based on this view, the user can choose to adapt her privacy settings, request removal of the information or report abuse to the social network.

To determine the visibility of a piece of information, it is necessary to consider all the users that should be involved in the management of that information, including the data subject(s). Tagging seems to be a viable solution for the identification of data subjects. Few approaches have been proposed to assist users in tagging images by identifying the subjects in an image. For instance, Stone et al. [24] propose a method for autotagging images within social networks which increases recognition performance beyond that of a baseline face recognition system. Facebook also provides an "autotagging" mechanism that automatically suggests possible tags for images. However, these tags often remain suggestions; moreover, they are only given to the data host who may not have incentive to accept them. In contrast, these tags should also be suggested to the data subject who can check the image and "accept" the tag. This approach has the added benefit of increasing data subject awareness which is a crucial step in improving user privacy in social networks.

To generate the complete visibility of a piece of information, the application would need access to the privacy settings of all involved users. However, in Facebook users can only access the settings of their own profile. Therefore, the proof-of-concept proposed in this paper cannot be deployed as a third party application. It should be provided by Facebook as a functionality of the social network (which has the necessary administrative rights) like the "view as" functionality.

6 Related Work

Privacy in social networks has been extensively researched [7,8,19,30]. Many research efforts have been devoted to the design and analysis of privacy enhancing systems for

social networks [14,20,31]. However, to the best of our knowledge, the impact of tagging on privacy and its relation with the multi-ownership problem have not been studied.

Several research efforts focus on the design of new privacy-friendly social networking sites. This is because existing social networks are often proprietary, and so it is difficult to validate the proposed approach. For instance, Cutillo and Mulvo [6] develop Safebook, a social network that handles privacy by real-life trust relationships. Baden et al. [3] propose Persona, a social network that uses user-defined privacy and attribute-based encryption. One attempt to enhance the privacy of existing social networks is Lockr [29]. Lockr is an access control system in which the social networking content is decoupled from functionality of social networks sites. This effectively removes the link between the user and the information and therefore enhances the control users have over their information. Differently from [29], the goal of this paper is to study the visibility of information in existing social networks rather than enhancing their privacy controls. In particular, our work is complementary to [29]: by having a complete view of the visibility of their information, users can effectively evaluate the consequences of sharing information and use existing techniques to restrict the access to it.

Another stream of research for enhancing privacy in social networks focuses on methods for controlled sharing of information. Controlled sharing among multiple users is often studied in the area of collaborative access control. Collaborative access control aims to balance the competing goals of collaboration and security [28]: collaborative systems aim to facilitate the sharing of information, while security aims to protect the same information. Tolone et al. [28] identifies the access control requirements for collaborative systems and analyzed existing authorization mechanisms with respect to such requirements. Although the identified requirements are general and applicable to social networks, they do not consider the new modalities of social communication that are emerging within social networks. Shen and Dewan [21] propose an access control model for collaborative editing. The model provides users a multi-dimensional, inheritance-based scheme for specifying access rights. Thomas [26] proposes team-based access control as an approach to applying role-based access control in collaborative environments. Yet, these models do not consider the new modalities of social communication. In addition, they mainly focus on organizational environments rather than on social networks.

In the context of social networks, a simple solution to the multi-ownership problem is proposed by Thomas et al. [25], in which the visibility of information is determined as the intersection of the privacy preferences of all involved users. This method, however, may result in the unavailability of information when, for instance, two or more users restrict the visibility to “only me”. Maximilien et al. [18] propose a privacy enhancing technology for evaluating profile privacy based on risk scores. Risk scores represent the level of privacy of a profile based on a comparison with other profiles and their respective settings. Recommendations are then made to lower the risk score based on the settings of other profiles. Another approach to collaborative access control in social networks is proposed in [23]. This approach maps user privacy specifications to an auction based on the Clarke-Tax mechanism [5] in order to select privacy policies that maximize social utility. In particular, privacy settings for an object are determined by

a collaborative decision between the involved parties. However, this approach requires the data host to identify all involved parties and provides no proper incentive to do so.

7 Conclusions

Social networks are offering their users new modalities of social communication for sharing information and building social relations. These new modalities, however, introduce new privacy issues. In this work we have investigated the impact of privacy settings and tagging on the visibility of information in Facebook. The analysis of scenarios representative for real situations has shown that privacy controls in Facebook are unintuitive. In particular, they may provide users a false confidence of being in full control of their information and therefore they may lead users to underestimate the risks of information sharing. To increase awareness and enable users to control the access to their information, we implemented Facebook's privacy controls to determine the actual visibility of information also in presence of tags. The main characteristic of the proof-of-concept tool is that it is information-centric (as opposite to object-centric) in the sense that visibility is determined by considering all the occurrences of a piece of information.

The work presented in this paper suggests some interesting directions for future work. The proposed tool implements Facebook's privacy controls as they are. Such controls, however, suffer from a number of drawbacks; e.g., their enforcement is not transparent for the user and do not consider the data subject as a main actor in the decision making process. Existing work on collaborative information sharing does not fully solve the privacy problems in social network. Indeed, most approaches do not consider the various functionalities like tagging available nowadays in social networks, leaving out crucial information about content visibility. We argue that new approaches for collaborative information sharing on social networks are needed. Such approaches should support the full range of functionalities for information sharing offered by social networks. Moreover, they should make users aware of the consequence of privacy controls they define and the risks of sharing information. For instance, they should provide transparent conflict detection mechanisms that assist users in the identification and resolution of conflicts with the privacy settings of other users. Social networks will therefore need tools that assure that data are accessed according to the user's settings, or notify the user why the settings of other users have been assigned a higher priority.

References

1. Apt, K.R.: Introduction to logic programming. In: The Handbook of Theoretical Computer Science. North Holland (1990) 495–574
2. Atwan, G., Lushing, E.: The Facebook Book. Abrams Image (2008)
3. Baden, R., Bender, A., Spring, N., Bhattacharjee, B., Starin, D.: Persona: an online social network with user-defined privacy. SIGCOMM Comp. Commun. Rev. **39**(4) (2009) 135–146
4. Bilge, L., Strufe, T., Balzarotti, D., Kirda, E.: All your contacts are belong to us: automated identity theft attacks on social networks. In: Proc. of WWW, ACM (2009) 551–560
5. Clarke, E.H.: Multipart pricing of public goods. Public Choice **11** (1971) 17–33

6. Cutillo, L.A., Molva, R., nen, M.: Safebook: a privacy preserving online social network leveraging on real-life trust. IEEE Communications Magazine (2009) 94–101
7. Faliagka, E., Tsakalidis, A., Vaikousi, D.: Teenagers' Use of Social Network Websites and Privacy Concerns: A Survey. In: Proc. of PCI, IEEE (2011) 207–211
8. Gross, R., Acquisti, A.: Information revelation and privacy in online social networks. In: Proc. of Workshop on Privacy in the Electronic Society, ACM (2005) 71–80
9. Guarda, P., Zannone, N.: Towards the development of privacy-aware systems. Information & Software Technology **51**(2) (2009) 337–350
10. Gürses, S.F., Rizk, R., Günther, O.: Privacy design in online social networks: Learning from privacy breaches and community feedback. In: Proc. of Int. Conf. on Information Systems, Association for Information Systems (2008) 90
11. Haron, H., Yusof, F.: Cyber stalking: The social impact of social networking technology. In: Proc. of Int. Conf. on Education and Management Technology, IEEE (2010) 237–241
12. Johnson, C.Y.: Project 'Gaydar' At MIT, an experiment identifies which students are gay, raising new questions about online privacy. Boston Globe (2009)
13. Larcom, G., Elbirt, A.: Gone Phishing. IEEE Technol. Soc. Mag. **25**(3) (2006) 52–55
14. Li, Q., Li, J., Wang, H., Ginjala, A.: Semantics-Enhanced Privacy Recommendation for Social Networking Sites. In: Proc. of TrustCom, IEEE (2011) 226–233
15. Luo, W., Xie, Q., Hengartner, U.: FaceCloak: An Architecture for User Privacy on Social Networking Sites. In: Proc. of CSE, IEEE (2009) 26–33
16. Mack, E.: Facebook's 'Like' button illegal in German state. CNET news (2011)
17. Marlow, C., Naaman, M., Boyd, D., Davis, M.: HT06, tagging paper, taxonomy, Flickr, academic article, to read. In: Proc. of Conf. on Hypertext and Hypermedia, ACM (2006) 31–40
18. Maximilien, E.M., Grandison, T., Liu, K., Sun, T., Richardson, D., Guo, S.: Enabling Privacy as a Fundamental Construct for Social Networks. In: Proc. of Int. Conf. on Computational Science and Engineering, IEEE (2009) 1015–1020
19. Nagle, F., Singh, L.: Can Friends Be Trusted? Exploring Privacy in Online Social Networks. In: Proc. of Int. Conf. on Advances in Social Network Analysis and Mining, IEEE (2009) 312–315
20. Qing-jiang, K., Xiao-hao, W., Jun, Z.: The (P, α, K) anonymity model for privacy protection of personal information in the social networks. In: Proc. of Int. Conf. on Information Technology and Artificial Intelligence, IEEE (2011) 420–423
21. Shen, H., Dewan, P.: Access control for collaborative environments. In: Proc. of Conf. on Computer-Supported Cooperative Work, ACM (1992) 51–58
22. Spiekermann, S., Cranor, L.: Engineering privacy. TSE **35**(1) (2009) 67–82
23. Squicciarini, A.C., Shehab, M., Wede, J.: Privacy policies for shared content in social network sites. The VLDB Journal **19**(6) (2010) 777–796
24. Stone, Z., Zickler, T., Darrell, T.: Autotagging Facebook: Social network context improves photo annotation. In: Proc. of Computer Vision and Pattern Recognition Workshops, IEEE (2008) 1–8
25. Thomas, K., Grier, C., Nicol, D.M.: unFriendly: multi-party privacy risks in social networks. In: Privacy Enhancing Technologies. LNCS 6205, Springer (2010) 236–252
26. Thomas, R.K.: TeaM-based access control (TMAC): a primitive for applying role-based access controls in collaborative environments. In: Proc. of Workshop on Role-Based Access Control, ACM (1997) 13–19
27. Thompson, H.H.: How I Stole Someone's Identity (Anatomy of a Social hack). Scientific American (2010)
28. Tolone, W., Ahn, G.J., Pai, T., Hong, S.P.: Access control in collaborative systems. ACM Comput. Surv. **37**(1) (2005) 29–41

29. Tootoonchian, A., Saroiu, S., Wolman, A., Ganjali, Y.: Lockr: Better Privacy for Social Networks. In: Proc. of Int. Conf. on Emerging Networking EXperiments and Technologies, ACM (2009) 169–180
30. Young, A.L., Quan-Haase, A.: Information revelation and internet privacy concerns on social network sites: a case study of facebook. In: Proc. of Int. Conf. on Communities and Technologies, ACM (2009) 265–274
31. Yuksel, A.S., Yuksel, M.E., Zaim, A.H.: An Approach for Protecting Privacy on Social Networks. In: Proc. of Int. Conf. on Systems and Networks Communications, IEEE (2010) 154–159
32. 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: Proc. of WWW, ACM (2009) 531–540