Audience Segregation in Social Network Sites

Abstract— In recent years research has shown that most social network sites pose serious privacy and security risks for individual users. From the existing analyses of privacy and security risks in social network sites we deduce that one of the biggest categories of privacy risks revolves around the notion of ‘audience segregation’, i.e. the partitioning of different audiences and the compartmentalization of social spheres. Since audience segregation is an important mechanism in everyday interactions between people in the real world, we argue that social network sites ought to include this mechanism as well. Current social network sites lack this mechanism. We present Clique, a privacy-preserving social network site that provides ‘audience segregation’ to its users as an alternative.

Keywords: social network sites, audience segregation, privacy, identity management

I. INTRODUCTION

Social software, ranging from forums, online communities, blogs, and dating sites to social network sites such as Facebook, LinkedIn and MySpace, has conquered the world. In this article we discuss some of the privacy issues surrounding the presentation of personal content (e.g., name, address etc.) and personal information (e.g., name, address etc.) in social network sites. Particularly, we examine users’ abilities to control who has access to the personal information and content they post in such communities. Social network sites lack a mechanism commonly used by individuals in their everyday interactions, that enables them to manage the impressions they leave on others and protect their privacy; audience segregation. We show that the mechanism is not only important in real life, but could also be a vital mechanism for the protection of one’s self-images and privacy in social network sites. In this paper we outline how audience segregation can be incorporated in social network sites, by presenting a prototype of a privacy-preserving social network site called Clique, which is under development as part of the EU FP7 PrimeLife project (http://primelife.eu). In Clique we have embedded three tools for audience segregation: contact management, setting visibility rights, and managing multiple faces in a single social network environment. We will discuss each in turn.

II. PRIVACY ISSUES IN SOCIAL NETWORK SITES

One of the fastest growing online fora for self-presentation and social interaction are ‘social network sites’ (SNSs). In June 2008 these sites attracted “an average of 165 million unique visitors a month” [1: 16]. In early 2010, Facebook alone claimed to have over 400 million users. In these online domains, users can present themselves using a so-called ‘profile’, and can interact with networks of ‘contacts’ within the same environment. Boyd and Ellison define social network sites as web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site [2: 211].

Sharing personal content and personal information is paramount to social network sites. Individuals join them to present information about themselves, for instance through text (e.g., blogs, status updates), through pictures, movies and sound clips, and through listing their ‘favorites’ regarding a broad category of pre-defined and user-generated topics ranging from clothing and brands, to music and movies, to locations and activities. This creates an image of each individual user. This image is primarily created by user himself, but other users may also contribute to it, thus enriching the projected image.

One of the most fascinating aspects of this new form of self-presentation is the degree of openness of most users [3, 4]. As Acquisti and Gross write: “...one cannot help but marvel at the nature, amount, and detail of the personal information some users provide, and ponder how informed this information sharing is” [5: 2]. Grimmmelmann dryly points out:

Facebook knows an immense amount about its users. A fully filled-out Facebook profile contains about 40 pieces of recognizably personal information, including name; birthday; political and religious views; online and offline contact information; sex, sexual preference and relationship status; favorite books, movies, and so on; educational and employment history; and, of course, picture. [...] Facebook then offers multiple tools for users to search out and add potential contacts. [...] By the time you’re done, Facebook has a reasonably comprehensive snapshot both of who you are and of who you know. [6: 9]
What makes people behave this way, given that there are obvious security and privacy issues? Why do users provide such detailed, and true, personal information on their social network site profile? Many explanations can be given, but we restrict ourselves to some of the most common ones. Acquisti and Gross list the following reasons: “Changing cultural trends, familiarity and confidence in digital technologies, and lack of exposure or memory of egregious misuses of personal data by others” [5: 2]. Grimmelmann argues that people misunderstand the risks involved in presenting detailed and personal information online. This misunderstanding takes a number of forms. For one, users are often unaware of who has access to their personal profile and to the content they place online, because the architecture and design of social network sites provide individuals with a false sense of security and privacy. These sites “systematically deliver them signals suggesting an intimate, confidential, and safe setting” [6: 17], an environment that is private, “closed to unwanted outsiders.” [6: 18]. Second, users falsely believe that there is safety in numbers, in two senses of the expression. They believe that when everyone else around them massively starts using social network sites, these sites must be safe to use, because otherwise others would avoid them (a line of reasoning that runs the obvious risk of being flawed if everyone follows it), and they believe the risks they run are limited since there are so many members in social network sites that chances are in fact really small that something will befall them as individuals [6: 17-18; 7: 133].

Taking things to a more general level, one can argue that there are four fundamental issues surrounding privacy and (unintended) information disclosure in relation to online worlds [8]. These can be summarized as follows:

I. One's audience usually is opaque when publishing information online;

II. Information published on the internet is persistent.

III. Information shared online may easily be transported to other contexts;

IV. Controlling self-presentations and the inferences based thereupon by others, is difficult for the individual.

These four issues are highly relevant to social network sites as well. For one, when posting content in a profile, individuals do not know (exactly) who will be able to access this information. Their audience, to phrase it differently, is not transparent. Although some social network sites allow users control over the visibility of their content (e.g., by making it ‘visible to friends only’), the default privacy settings are usually set to ‘public’, which means that it can be viewed by anyone accessing the social network site.

Second, since information can be copied easily, it can be re-published at any particular moment and may come back to haunt the individual years down the line. As a result one’s audience is unlimited both in terms of its size and makeup (in contrast to audiences in the physical world), but also in terms of temporality. The primary audience of information may (unintentionally) exist in the future [4: 22].

Third, when presenting disparate identities in various online domains, there is a risk of information spilling from one context, for instance a home page, into another, such as one’s social network site profile. Since different behavioral rules guide these various domains, mixing and merging information about an individual can lead to serious problems. Tufekci illustrates this nicely: “For example, a person may act in a way that is appropriate at a friend’s birthday party, but the photograph taken by someone with a cell phone camera and uploaded to MySpace is not appropriate for a job interview, nor is it necessarily representative of that person. Yet that picture and that job interview may now intersect.” [4: 22]

Last, in social network sites who we are is expressed by an online representation of ourselves, which may be composed of, for instance, a profile with personal details, stories and pictures. Our control over the type and content of information we put online only goes so far. Other users can add information to our profile or alter it, put pictures or information about us on their own or other people’s profiles, and tag pictures to reveal the identities of those portrayed in them. Tufekci’s example is a case in point: placing a picture of another person online affects the image of that person to the audience viewing it, and hence may have an effect on the (current and future) self-presentations and impressions of that individual.

The central question we posed ourselves is whether we can contribute to solving some of the issues outlined above.

III. The Importance of Audience Segregation

In our view, there are two central issues to be addressed relating to privacy issues in social network environments:

I. User awareness of the privacy issues discussed above should be raised. Users need to become more aware of the fact that and the ways in which information may ‘leak’ to unintended audiences on the internet;

II. Users should be provided with tools to help them manage their personal information and content in a more privacy-friendly manner.

To maximize awareness and usability, these tools ought to be easily recognizable for users. This is why we turned to a mechanism that individuals use in everyday life contexts to control the image others have of them and the information they disclose about themselves: ‘audience segregation’. Mirroring or mimicking this real-life strategy, we have developed a social network site, Clique, in which a number of instantiations of audience segregation are implemented.

A. Audience segregation: theoretical background

The concept of ‘audience segregation’ was coined by Erving Goffman [9] as part of a perspective on the ways in which identities are constructed and expressed in interactions between human beings in everyday contexts. According to Goffman, whenever individuals interact with others, they perform roles, with which they hope to present a favorable image of themselves. To Goffman, impression management is key in such self-presentations.

Each person performs a wide variety of roles in his everyday life, relating to both the places they visit, and the other people
Individuals attempt to present self-images that are both consistent and coherent. To accomplish this, they engage in 'audience segregation', "...so that the individuals who witness [them] in one of [their] roles will not be the individuals who witness [them] in another of [their] roles" [9: 137]. With segregated audiences for the presentation of specific roles, people can 'maintain face' before each of these audiences. Their image will not be contaminated by information from other roles performed in other situations before other audiences, particularly not by information that may discredit a convincing performance in the current situation [9: 137]. For example, a person whose professional role consists of displaying authority, such as a political leader, may try to shield not being in charge at all when at home. Shielding this fact from those encountered in professional life helps him to maintain his professional authority.

Audience segregation and privacy are closely linked. Helen Nissenbaum argues that privacy revolves around 'contextual integrity'. She writes:

> Observing the texture of people’s lives, we find them [...] moving about, into, and out of a plurality of distinct realms. They are at home with families, they go to work, they seek medical care, visit friends, consult with psychiatrists, talk with lawyers, go to the bank, attend religious services, vote, shop, and more. Each of these spheres, realms, or contexts involves, indeed may even be defined by, a distinct set of norms, which governs its various aspects such as roles, expectations, actions, and practices. [13: 137]

Nissenbaum argues that privacy means respecting the contextual boundedness of the (personal) information individuals share in each of these distinct realms. Phrased differently, according to this perspective privacy revolves around individuals’ ability to keep audiences separate and to compartmentalize their (social) life.

**B. Audience segregation in social network sites: why?**

Above we have argued that users in social network sites lack mechanisms to separate and manage the various audiences for whom they perform. Many social network sites cluster all of an individual’s contacts into a single category, called ‘friends’. Given the fact that the average Facebook user has 140 ‘friends’, this necessarily conflates different contexts. This means that a) it is impossible for users to hide parts of their network of contacts from other contacts (e.g., I do not want my colleagues to see my friends, or I do not want my mother to see my colleagues); and b) that it is impossible to restrict access to information to part of their network.

Providing users with mechanisms to control access over the information they present in social network sites would improve the quality of interactions and self-presentation. First, it would mimic real life interaction patterns to a larger degree, and align more closely with the ways in which individuals tend to engage with others in everyday settings. Second, enabling better access control and audience segregation in social network sites could effectively counter some of the privacy and security risks we have discussed above and, therefore, make social network sites more privacy-preserving [also see 14]. Given the numbers of people active on these sites today this is a worthwhile goal to strive for indeed. Third, enabling users to compartmentalize the audiences for whom they perform in social network sites allows them to present different sides of themselves to different audiences, thereby allowing each (partial!) self-presentation to be textured and full of depth. This will help users avoid what boyd calls ‘social convergence’, the presentation of a single face that is acceptable to people that belong to different audiences [15].

Social convergence occurs when disparate social contexts are collapsed into one. Even in public settings, people are accustomed to maintaining discrete social contexts separated by space. How one behaves is typically dependent on the norms in a given social context. [...] Social convergence requires people to handle disparate audiences simultaneously without a social script. While social convergence allows information to be spread more efficiently, this is not always what people desire. As with other forms of convergence, control is lost with social convergence. [15: 18]

Audience segregation allows users to be ‘round characters’ in different roles, rather than ‘flat ones’ in a conflated context.

In many social network sites, including Facebook, Friendster and MySpace, individuals currently connect with both friends, family members, distant relatives, colleagues, acquaintances, old schoolmates, members of their local community, etc. – some of whom are intimately known to them, while others are distant, loose, or even unknown connections. It is easy to see why individuals using such sites might want to make distinctions between the types of information they want to make available to each of these different categories of connections, and give different connections access to different content. For instance, a user might want to share his holiday pictures with close friends, family members and other relatives, but not with colleagues or old schoolmates. Or, more specifically, he might want to share his holiday pictures with his close friends and family members – but not with Mom and Aunt So-and-so.

Currently most SNSs provide limited options for making one’s profile or its content (in)visible for specific others or specific collections of others. Generally, users can choose from: ‘visible to everyone’ (i.e. all members of the social network site), ‘visible only to friends’ (i.e. all of the user’s contacts!), ‘visible only to friends and friends of friends’, and in some cases ‘invisible to everyone’. On some sites, users can specify the (in)visibility settings of specific types of information, e.g., they can make their basic information (name, home town etc.) available to all members of the site, while restricting access to their pictures to their contacts. Assigning different ‘collections’ within one’s own network of contacts has recently become available in some networks, such as Facebook, but it is very difficult to use in practice.
C. Terminology used in Clique

The language used to discuss social network sites, the users participating in them, and the connections between these users is often quite fuzzy and imprecise. This is why we pause to define each of the concepts we have used in Clique.

I. The terms ‘platform’ and ‘social network site’ will be used interchangeably;

II. On the platform a user can create a ‘face’, a profile page to present particular information about him self. The totality of all the faces a person manages within a platform makes up his identity. Currently, most platforms allow users to create just one face;

III. ‘Contacts’ are all the individuals with whom a user is connected within the platform;

IV. ‘Collections’ are sets of contacts selected by the individual from the totality of his contacts. The user can assign a name to each collection to identify them (e.g., ‘best friends’, ‘colleagues’, etc).

V. A ‘context’ is the combination of a particular face and its associated collections. For instance, a ‘work context’ is one in which a user presents his ‘work identity’ (face) to his ‘colleagues’ and ‘former colleagues’ (collections).

We have summarized this terminology in figure 1.

IV. FROM CONCEPTS TO PRACTICAL TOOLS

In the remainder of this paper we present our implementation for realizing audience segregation within Clique, using three tools: a tool for contact management, one for setting access control policies, and one for managing multiple faces.

A. Contact management: collections

Audience segregation is based on nuances in connections [also see 16: 72]. This means that users are able to create their own social clusters (collections), in which they group one or more of their contacts, and that they can assign labels to these clusters. This departs from most current-day social network sites, in which all contacts in a user’s network are lumped together in one collection of ‘friends’. By allowing users to create collections within their list of contacts, they can cluster social relations according to their own preferences, thereby mimicking the actual practice of building and maintaining separate social spheres in real life. Users must be free to define (and label) their own collections, since that is the only way in which these collections will correspond to the fabric of their social life. Grimmelmann [6] has argued that if the provider of the social network site offers the possibility to place contacts in clusters (such as ‘family’ or ‘friends’), these clusters could never be an adequate representation of the complexity of social relationships in real life. He writes:

Consider the RELATIONSHIP project, which aims to provide a ‘vocabulary for describing relationships between people’ using thirty-three terms such as ‘apprenticeTo,’ ‘antagonistOf,’ ‘knowsByReputation,’ ‘lostContactWith,’ and ‘wouldLikeToKnow.’ [...] Clay Shirky shows what’s wrong with the entire enterprise by pointing out that RELATIONSHIP’s authors left out ‘closePersonalFriendOf,’ ‘usedToSleepWith,’ ‘friendYouDontLike,’ and every other phrase we could use to describe our real, lived relationships. […] We shouldn’t expect Facebook’s formal descriptors to be precise approximations to the social phenomena they represent. [6: 27]

Grimmelmann is correct in claiming that the platform provider cannot capture the complexity of individuals’ many social spheres and connections. However, we argue that the individuals themselves are fully capable of doing so. We all know which individuals make up our social circle and what the different degrees of intimacy in that social circle consist of. In Clique, therefore, we have built a tool for contact management that allows users to replicate their social sphere in any level of granularity that works for them. This solves the problem signaled by Grimmelmann above.

In Clique users can cluster contacts into self-assigned and self-labeled sets. After inviting contacts, they are asked to assign them to one or more ‘collections’, which can be changed at any time. Figure 2 shows collection management in Clique. Note that the collection ‘colleagues’ is marked as Ronald’s primary audience (marked as default).
B. Setting visibility rights

The second principle in realizing audience segregation in social network sites relates to contextualizing the user’s profile and all the information published there [16: 72]. This means that information is made public for a specific audience, which may be made up of one or more collections, and/or one or more separate individuals. In Clique, contextualizing content and information is implemented by means of two tools. The first is discussed in this section; the second is discussed below. The first mechanism is the use of visibility rights, which enables users to assign access rights to different collections and individuals. Each time users post items of information or content within a context, they can choose for which audience (both collections and individuals) this item will be visible. For example, a user may decide to make his holiday pictures invisible to his colleagues, but visible to his relatives and some members of his collection of friends, or he may decide to prevent acquaintances from reading his diary entries, but leave it visible to everyone else in his contacts list.

In Clique, individual users can control visibility settings of each individual item of information for two reasons. First, individuals use social network sites to present content with different goals and purposes in mind. Some may use these sites, for instance, only to stay in touch with people they know intimately in the offline world, whereas others may want to use them especially to present (aspects of) themselves before an audience of strangers. Obviously, users thus have different requirements regarding the visibility of their information. Therefore, it would be patronizing and limiting if the platform provider would decide for users which information to share and for which (limited or unlimited) audience. Second, users’ ideas of which kinds of information are deemed ‘private’ vary: “Different people have different views of what should be private. [...] People must be able to reach their own decisions about what should be private, and what gains they would hope to make by releasing information about themselves” [17: 74].

An objection to providing extensive control over visibility settings could be that users don’t want too much control over their content in social network sites. In fact, researchers have argued that users are not interested in fine-grained control over the display of personal data, for instance because making the profile invisible makes it harder for other people to find them [7], or because they would simply find it too much hassle. However, recent research has shown that, when given the opportunity, many people do in fact want to shield some of their information [3], especially since quite a significant number of negative examples regarding information spill and privacy issues with respect to social network sites have been published in the press in recent times.

Clique implements a fine-grained mechanism for setting access control policies, in which each element of the profile can be made visible for either collections, or individuals, or a mixture of both. This means, for instance, that a user can make his name and date of birth visible to everyone, while restricting access to his address to colleagues, and allowing only some designated contacts to see his mobile phone number. Figure 3 shows a user profile page in Clique. Each item contains an icon that displays its current audience on mouse over (see figure 5). Users can choose between the following access control options for the content published on their profile: ‘only visible to me’, ‘contacts/collections’ (e.g., mobile phone), ‘all contacts’ (e.g., website), and ‘public’ (e.g., location).

Figure 3: Visibility settings in Clique.

When users publish information in Clique they are presented with an access control dialog as shown in figure 4. In this dialogue window we ‘nudge’ [18] the user to act in a privacy-savvy manner without undermining sociality. By default, the user’s primary audience (default collection, see figure 2) is selected as having access to the content to be published. The user can drag collections and individual contacts to the red and green boxes to grow or shrink the audience. Note that in this case, Ronald’s colleagues have access to the content to be published, with the exception of a few contacts. While enabling access to a collection, thus, the user can still choose to make information unavailable for particular individuals.

Figure 4: Extended access control dialogue in Clique.

C. Managing multiple faces on a platform: tabs

The third tool we have developed is the introduction of tabs to represent the different ‘faces’ a user may want to combine within the same platform. This form of contextualization mimics the fact that individuals maintain different social spheres in the offline world. Most social network sites implement a single profile for each user. All of a user’s contacts see the same in-
formation. As we have argued, it is important to allow users to diversify the information and content they present to different audiences. Moreover, many people now maintain different profiles on different platforms, which is cumbersome and time-intensive. If these profiles could be combined in a single social network site, users would only have to access that single environment to manage multiple, separate self-presentations.

In Clique, the different ‘faces’ a person may have in the offline world can be recreated using tabs. Each tab functions as a separate social sphere, representing one aspect of the user’s identity. For instance, users may create a tab for their private face and for their professional face. Each of these faces contains its own network of contacts, which can be assigned to the various collections within each tab. Access rights can be defined for collections and contacts with regard to all content presented in this context (i.e. using a specific face in front of a specific collection). Contacts only get access to the information that is made visible for them. This means that a) contacts who only know the individual professionally, for instance, are prevented from acquainting themselves with the user’s leisurely profile; and b) within each face, contacts can only access the information that is explicitly made available to them.

The tabs (see figure 3) to distinguish between different contexts are a visually appealing and easy way for the individual to manage their various profile pages (faces) in Clique. Information added to one of the faces (e.g., the ‘Ronald’ tab) is invisible in all other tabs, and hence it is easy for the user to manage who sees what. Clique can therefore be used as a dashboard for multiple social environments. By simply clicking through the different tabs the user can see what information is accessible there, while the audience indicator icons reveal the current audience.

Creating faces is a bit cumbersome, since it means that users need to build a new profile, set the security and privacy settings, and add contacts and content for each individual face. They have to invest energy and time in setting up a new profile. Particularly when users create multiple faces for which the contact list shows a significant overlap we may wonder whether users are willing to make this investment, and whether they may see (enough of) the benefits and advantages of creating separate faces.

IV. Conclusion

Context is a central concept in the disclosure of information. What is appropriate in one context is not in another. We have argued that audience segregation is one of the core mechanisms

that people employ in their everyday life to accomplish contextual integrity and that most current online social network sites have a very simplistic model of social structures. In our view, technology can be adopted to help users maintain different partial identities en control who can access their data even in social networks. Whether or not social network site users can and will use the mechanisms provided remains to be seen. To test whether they do, we have set up an experimental site consisting of the Clique prototype (http://clique.primelife.eu). The reader is invited to participate in this experiment.

References

(6) Grimmelmenn, J.: ‘Facebook and the social dynamics of privacy [draft version]’, (2008), pp. 1-52