Privacy issues in social network sites

Social network sites are among the most popular web 2.0 domains worldwide. Millions of users have a 'profile page' on Facebook, LinkedIn, MySpace etc. In early 2011 Facebook alone has had more than 500 million active users. As social network sites proliferate, so does media coverage and scientific research on the privacy issues their use may generate. Users tend to disclose highly personal, detailed information via these sites, and to share this information with large groups of (vaguely known) 'friends'.

Four issues relating to privacy can be distinguished in literature on social network sites:

1. When posting content or personal information on a profile page, individuals do not know (exactly) who will be able to access this information. The audience is intransparent.

2. Since information can be copied and saved easily and indefinitely, information placed online at any particular moment may come back to haunt the individual years down the line. The audience is unlimited not only in its size and makeup, but also in a temporal sense.

3. In the current situation, many users have profile pages on multiple social network sites, to display different aspects for their identities to different audiences. For example, they maintain a professional profile on LinkedIn, and a more personal one on Facebook. However, information from one of these domains may very easily seep into another, and may thus 'contaminate' the partial identity displayed there.

4. In social network sites the amount of control users have over their self-presentation is limited. Other users can add or change information in a person's profile, share pictures or information about the person, and tag pictures to reveal the identities of those portrayed in them.

Building solutions for these privacy issues in social network site

In PrimeLife we asked ourselves: How can we design social network sites to optimise users' options for protecting their privacy and overcome some of the issues cited above? Our goal was to design a social network site that mimics the commonly used social practice of 'audience segregation' in the online world.
Audience segregation: Compartamentalising social spheres

The term 'audience segregation' was coined by the Canadian sociologist Erving Goffman in the 1950s to denote a common social practice: Individuals operate in different social settings as they go through their daily routines, and in each of these settings they show different sides of themselves. For example, they behave differently when at home than when at work, and again differently in the checkout line of a supermarket than at a bar with their friends. Audience segregation refers to the fact that individuals strive to leave a favourable impression in each of these settings, and hence would find it uncomfortable if information from one setting (drinking stories from the bar) spilled over to another one (a meeting with superiors at work). In short, audience segregation ensures that (under-mining) information from one context does not discredit one's performance in another.

While audience segregation is a commonly used social practice in the offline world, no virtual or online version of this mechanism has existed yet. This is puzzling since this single mechanism could solve several of the privacy issues in social network sites which we have discussed above.

Clique: Implementing audience segregation

Using Elgg open source software, we have built a social network called Clique, in which we have implemented three mechanisms for the creation of online audience segregation:

1. Contact management: Collections

In most current social network sites all of a person's contacts are lumped together into a single category called 'friends'. No distinction is made between the different social relations a person has (friends, family, acquaintances, colleagues, etc.) and information is made available to all 'friends', regardless of their relationship with the user. In Clique, users can cluster contacts into different 'collections'; which they can label themselves ('close friends', 'former classmates', 'neighbours', 'former colleagues' etc.). With each item of content users post, they can specify exactly which collection and/or individuals can access the content.

2. Setting visibility rights

In most current social network sites users cannot choose to change the visibility settings of specific items of personal information, e.g. their home address, phone number, or gender. In Clique, users can set visibility rights to each individual item of information. Small icons accompanying the specific item of information provide feedback on the chosen setting, and a mouse-over leads to detailed information regarding the size and makeup of the selected audience.

3. Multiple faces

To ensure that users can display different sides of themselves to different audiences, we have implemented 'faces' in Clique. Each face is represented by a different tab on the profile page, and different contacts, different collections and content may be added to each. Using tabs is a visually appealing way to raise user awareness relating to the contextual nature of self-presentation in social network sites.

Scramble your Clique content

In most social networks, the providers have access to all of the users' information, public and private. To protect the content data from unwanted access by the social network providers, PrimeLife provides the tool 'Scramble!' that works nicely together with Clique.

Scramble! is a tool that allows not only for the definition of access control rules for audience segregation, but also for the support of their enforcement. It allows users to publish encrypted information and define the set of users that have the ability to decrypt – and therefore view – the published content.

Scramble! is a Firefox extension. It is independent of the actual setup of social network sites. Once set up with a user's cryptographic key pair and a suitable set of users and groups, Scramble! will automatically look for content, encrypted with the public key of the user. Such information blocks are parsed, decrypted and represented in the browser without any user interaction. Generating encrypted content is done by selecting the data, clicking one button, and going through a selection dialogue to construct the intended audience. Public keys are stored online in an OpenPGP key server and can be retrieved easily. Finally, users that have no access to the encrypted data will see a suitable message (non-authorised content) while users that do not have the extension installed see a strange-looking but innocent piece of data.