Privacy in Online Social Networks

Ahmad Mushtaq
Helsinki University of Technology
ahmad.mushtaq@hut.fi

Abstract

Social networking is the latest craze that has captured the attention of masses, people use them mainly to communicate with their real life friends, but these sites claim to be your virtual life on the internet where one could do most of the things they do in real life. Privacy is one issue that suddenly comes to mind when your information is available online for every one to see. In this paper we examine how privacy of a user is affected by social networking sites, what kind of information users share with in their online profiles, what can be the implications for users when sharing information. Before ending the paper I will also discuss some of the steps that can be taken to help protecting one's privacy when using such social networking web sites.

KEYWORDS: privacy, social networking, online privacy facebook, orkut, linkedin, myspace

1 Introduction

In recent years, online social networks has become very popular, many web sites have sprung up where one can meet their offline friends in virtual world of the internet. Services like Facebook, Orkut, MySpace etc allow people to host their online social networks, people create their profiles in such social networks and share this information with their friends and a vast amount of strangers on these social sites.

1.1 Profiles in online social networks

Most online social networking web sites revolve around a similar core feature: "user profiles" or just "profiles". Profiles are an important element of these social networks, and a users profile contains an array of information about the user, describing himself with elements such as physical appearence, hobbies, personal photo and/or pictures of friends, contact information and a vast array of other user contributed content.

1.2 Sharing Personal Information

When people share their personal information over a social network, it is not clear who can read the information, can every one read that information? or only the people from the friend list of the user see that information? but also and more importantly how does this affect the privacy of the user.

Social networking websites are now coming up with new innovative ideas for the users of the sites, Facebook for example now lets users to install applications in their profiles, these applications range from calendars to movie recommendation and review applications, dating, games and so on. Orkut is another example, Orkut is a social networking service owned by Google and thus combines the other services offered by Google, for example users can share their YouTube videos, their Picasa Photo albums, blogs among other things on their profiles. With profiles in such social networking web sites users are already sharing so much information, but when a social network combines the resources from other web sites, the privacy of the users can be exposed in a whole new manner.

1.3 Risks in sharing information

Thus in online cocial networks privacy can become a serious concern, when people share more information then just their name or age, information like their home address, their pictures and other sensitive information they can expose them selves to various risks. What can be those risks how exposing certain information can lead to those risks, and what can be the steps that can be taken to safeguard against such risks.

2 Structure and Evolution of Online Social Networks

The study of social networks has received signicant interest from researchers in various domains such as psychology, philosophy, education, and lately computer science - particularly in the field of artificial intelligence. I try to define what we mean by social networks, the way in which these networks form and evolve.

Social networks are a social structure of nodes that represent individuals and the relationships between them within a certain area. Therefore, social networks are usually built based on the strength of relationships and trust between the members. Scrutiny of the ways in which these nodes are connected has resulted in the identification of varying types of ties between nodes. In this context, a strong tie is one established directly between two people in the same network, whereas a weak tie is a relationship between two people connected through another person. [12] [4]

The way in which people meet and form social networks in everyday life has caught the attention of many computer science researchers. The fact that we relate to and depend on our social network for such things as friendship, support, special interests and knowl- edge sharing has inspired researchers to analyze these facts.

The way the social networking has evolved can be lined with the evolution of the internet itself.

Bulletin Boards

People have been using computers for social communication since the very beginning of the personal computer industry. Long before the Internet became accessible to the general public, people were hosting BBS systems, many of them focused on an interest group or local community, such forums had also had profiles, but unlike the online social networks of today they are more limited and usually contain harmless information.

Online Services

Commercial online services reached their peak in the 1990s, first as destinations in of themselves, and later as a way to access the Internet. These services provided access to a broad range of services that are now mirrored on the web. News, travel reservations, shopping and social hubs were all part of the package; much of what we see today on the web existed in some form on these sites. Social communication was one of the big draws for online services, as a major source of their revenue was derived from billing for usage on a per-minute basis. AOL in particular recognized this and allowed users to create communities about just about any topic.

Web 1.0

From the mid-90s to 2000, there was an explosion of activity as companies rushed to reproduce existing online services on the web. There were many social services created during this period, notably GeoCities and theGlobe.com. One thing the web did was to eliminate the walled garden problem that plagued AOL and their brethren. This promoted the development of niche communities, during this period privacy was not a major concern, internet was just taking off and services were given more importance then privacy, however the concept of social networking had not evolved enough yet, and people were in full control of their content and what they were posting online, but the absence of profiles limited privacy related risks.

Web 2.0, Profiles and Updates

Social networks make it easy for people to create profiles using standard templates. This makes sense, but this is really no different than a web page. People can share information about themselves, profiles can make it wasy for users to share too much information, and even if they dont share too much, the little that they share can also can impact on their privacy, since there are numerous social networks and information may be distributed over all of them even if little on every site.

One of the reasons Facebook is so addictive is because it is a convenient way to track the status of friends. This, too, is something that can be moved onto the open web. Anyone who wants to can publish updates, events, etc. via standard formats like RSS and iCal. Anyone who wants to monitor their friend's updates can do so, via a feed reader or some other way.

3 Privacy Revelation

Social networking sites share the basic purpose of online interaction and communication, but the specific goals and usage patterns change across different services. The most common scenario is based on the use of the participants profile and the presentation of her/his network of friends. This approach can stretch towards different directions. In dating web sites, like Match.com or Nerve and Salon Personals, the profile is the critical component and there is no network of friends. In blogging sites like LiveJournal, Blogger etc profiles become secondary, networks may or may not be visible, blog entries take a central role.

Online social networking thus can change into classifieds in one direction and blogging in another direction. The way people share information on these online social networks is very different and depends on the site under consideration. Using a real name to present an account profile to the members of an online community may be encouraged on sites like facebook which focus on college students[14], such networks try to connect participants profile to their real life identities. Other web sites like friendster tolerate use of real name, such sites create a thing layer of anonymity between the real life and online profile by making only a part of the real name visible on the online profiles, i.e. showing only the first name while hiding the last name. Other web sites like match.com discourage publishing real names and other personal contact information, such sites attempt to protect the real life identity of a person by making its link to the online profile more difficult.

Most social networking sites encourage the publication of personal and identifiable personal photos. The type of information revealed or found often revolves around different hobbies and interests, but can go to may other different directions, like information such as current and previous schools and employers; private information such as drinking and drug habits and sexual and political preferences and orientation, relationship infomation, for example who is the girl friend / boy friend, or spouse etc.

The visibility of information on online social networks changes all the time. In some sites any member may view any other member'ss profile. On other social networking sites access to personal information may be restricted to people that are part of the direct or extended network of the profile's owner. However the visibility control of the information changes across different sites. People generally are happy to disclose as much information as possible to as many people as possible[6].

Different fields used in online profiles and percentage of users filling them [5].

1. 93.8% of users disclose their sex.
2. 83.3% of users disclose their hometown.
3. 87.1% of users disclose their highschool.
4. 45.1% of users disclose their home address.
5. 59.8% of users filled out the About Me field.
6. 67.8% of users disclose their Instant Messaging address.
7. 83.8% of users disclose their birthday.
8. 92.3% of users disclose their email address.
9. 78.5% of users disclose their relationship status.

According to a report published by PEW Internet & American life project [10]:

1. 55% of online teens have online profiles, 45% of the online teens don't have profiles online
2. 82% of profile creators have included their first name in their profiles
3. 66% have included photos of their friends
4. 61% have included the name of their city or town.
5. 49% have included the name of their school.
6. 40% have included their instant message screen name.
7. 40% have streamed audio to their profile.
8. 39% have linked to their blog.
9. 29% have included their email address.
10. 29% have included their last names.
11. 29% have included videos.
12. 2% have included their cell phone numbers.
13. 6% of online teens and 11% of profile-owning teens post their first and last names on publicly-accessible profiles;
14. 3% of online teens and 5% of profile-owning teens disclose their full names, photos of themselves and the town where they live in publicly-viewable profiles.

4 Privacy Implications

The privacy implications of a social networking web site depends on the amount of information it provides, who can access that information and how accurate that information is, even if some social networking site protects the identity of a user, it may still be possible to link an online profile with the real world identity of the person, for example through face identification [13]. Often people have accounts on many social networking sites, and usually they use same picture across different sites, and on some sites they may provide more information in their profile then other sites. So through face identification, if a link between different profiles can be constructed, then it may be possible to gather more information, just it would be possible from just one social network. Identification can also be archived even if the profile picture information is not used, in such cases demographic information can be used, or though other overlapping information like special hobbies, interests etc, though this approach can be less fruitful.

Recently some social networks have started to expose their API's and have made it possible for developers to implement applications one famous example is Facebook. Many different kinds of applications are available, users can play movie trivia, where they can answer questions and they will be ranked with their friends, they can also be ranked how closely their movie tastes match with their friends. Then there are some applications where users can mark different cities or countries they have visited, have lived in or want to visit in the future. These applications can also help in the revelation of different kind of information about a particular user. Usually when a user installs some application or performs some activity with some application his network of friends are also notified about the activity, the friends are also invited to take part in the activity usually they can comment and share views about such activities.

How can this information be used, once some third person has access to a social network, risks range from identity theft to online may be even physical stalking, black mailing, embarrassment or other people getting to know about information not meant to be shared with them. People can get email address, ids of different messaging services used by the person, this information can then later on be used to stalk the person, female users are specially the target of such kinds of stalking. The profile pictures on display can also be misused by persons with negative agendas, in some cases these are just forwarded to friends, but in other extreme cases these may be posted on inappropriate forums or web sites. There have been some reported cases where people also called the person on their phone numbers to harass them, most of these phone calls are harmless but some times it can get serious.

The information shared on the social networking web sites is always available to the operating staff of the web site, and also to the web hosting provider of that online social network. Such information may be stored indefinitely when the site data is backed up even long after the user has removed his account from the online social network. This information will also be shared by either the social network provider or the web hosting provider with the law enforcement agencies upon their request.

4.1 Addiction to social networking sites

Social networking sites can be addictive[14]. If a user becomes addicted to such sites, then they care less about the amount and the kind of information they are posting online. Usually teenagers get addicted to such sites and may pose a privacy risk to the whole family when post information and not understanding the consequences that occur because of it.

4.2 Stakling

Stalking can be one real side-effect of using online social networks when not enough attention is paid when sharing information, in social networks that try to mirror real life communities into social networks such college or work place communities, it becomes easier to predict where a person could be in a specified period in time. For example a student is usually at hostel or attending some classes that he has enrolled into, by gaining such information it can be possible to determine his whereabouts. [11]

Users generally also share their instant messaging accounts, which may include Yahoo! Messenger, MSN Messenger, Google Talk, AOL Instant Messenger among others. Some of these services allow users to add friends or buddies to their list with out the knowledge or first getting a confirmation from the person being added to the list. Once the person is in the buddy list, it is possible to send messages or track the online presence and in some services its even possible to send messages.

4.3 Identification

In an experimented conducted by (information revelation and privacy in online social networks) it was possible to correctly match the pictures published in profiles by students in their online social network profiles with the pictures on the universities student pages using a commercial face recognition software. Usually the profiles contain high quality images of the person, which provide more than enough information for the face recognition algorithms to perform correct identification. According to [5] about 90.8 of all profiles on Facebook.com have some kind of pictures in their profiles, and out of this 61 of the profiles are estimated to have pictures of sufficient quality to be used for identification.

4.4 Identity Theft

By exposing such information as birth date, home town, residence and phone number users can be vulnerable to identity theft. This information can be used to estimate the social security number of a person, which can then lead to identity theft. However this approach has been shown to work in United States of America by [5].

4.5 Building a digital dossier

When young users build their online profiles, they generally are not concerned about their privacy and publish more personal and sensitive information. Given that the storage costs are declining, it is possible for governments to continuously harvest information and store, later on when these young people enter the professional life which may have sensitive and delicate situations, such information then might become important and can be used in situations where it can be harmful for their public image or relations[3][7] [8].

4.6 Manupulating users

Social engineering methods can also be used in online social networks, it can be used to retrieve more information about a users profile, just by simply asking them, according to [6] by using an automated script about off 250,000 users who were asked to accept a friend ship request from a facebook user using the script 75,000 users accepted the request and exposed their full profiles, thus it was deduced that 30 of all Facebook users may be willing to share their private information with a complete stranger.

5 Protecting Privacy in Social Networks

Online communities require you to provide personal information. They ask for at least a user name, e-mail address, and password when registering with these services. In the interest of community building and commercial marketing, later on the social networks might request to fill out a user profile that includes much more personal information, such as birth date, your home and work addresses, home and work phone numbers, gender, marital status, occupation, instant messaging names, and more. These profiles are usually public, but some social networking sites allow their members to view profile of any other members even if they are not in their social network of friends. Its usually a good idea not to post information that a person is not comfortable sharing with strangers. The comments posted are permanently recorded on the social networking site. As time goes by and users get to know other members of the community, the community might begin to feel casual and familiar, and users might be tempted to talk about their kids by name, mention where you work or live, or reveal information about valuable collections in your home that might not be listed in the profile or the user might have refrained from publishing in the profile. As the discussions get casual, users sometimes may even mention when they plan to be out of town. Its usually a good idea to stay vigilant at all times even if a member has been involved with a community for a long time, members should not be fooled into a false sense of security.

5.1 What users can do to protect privacy

Some information is so revealing of individual identity that it should never posted on the public area of any website, for example social security number, phone numbers, home address. Social networking sites like Facebook, MySpace, LinkedIn and similar sites have become incredibly popular because of their social factor: they help people stay in touch with friends or business contacts, and make new ones as well. To take advantage of their benefits, though, these services need people to share some personal information. That's where one should be cautious before sharing any information, people should understand the information that they share with the site, think about the information the site will share and with whom. Use the privacy controls the site provides to better protect information. It is also important to know that how the social networking site will use information themselves.

Users should refrain from posting sensitive information on social networks, things about their intimate life or work, things about their employers, even if these are well protected by the sites security mechanisms, there is always a danger of a hacker somehow getting access to the social networks server or just the users account[2]. Such hacker may gain access to the private information and might post it on public places, where it may stay for a long period of time if a search engine like Google keeps a copy of said records in its cache.

Online communities offer several ways to ensure privacy, before joing a social networking web sites, some of the things to look for are:

- Privacy policies that explain exactly what information the service will collect and how it might be used.

- User guidelines that outline a basic code of conduct for users on their sites. Sites have the option to penalize reported violators with account suspension or termination.

- Special provisions for children and their parents, such as family-friendly options geared towards protecting children under a certain age.
- Password protection to help keep the user account secure.
- Email address should be hidden on the profile, or atleast the social networking site should offer an option to the users to hide or show their email address and who can see their email address, options should be like so that user can select if only friends can see their email, or the friends of friends can see the email, if every one can see the email or if no one can see the email address.
- E-mail address masking: Masking involves inserting a word or phrase in the middle of your regular e-mail address to help foil automated e-mail "harvesting" programs, for example: someone@nospam.example.com. However, as spammers become more sophisticated, their harvesting software might be able to recognize a masked e-mail address.

5.2 What technology can do to protect privacy

Social networks can make it easier for users to use privacy controls, so that even a new novice user can control what other people can see from their profile. When a user creates a new account on a social network the default settings should be so that the private profile entries are hidden or even the whole profile is hidden unless the user chooses to publish his profile to public or his friends later on. Warnings should be display and online easy to understand help should be available whenever user chooses to change any of the privacy settings.

CAPTCHAs[1] can be used by the social networking sites to prevent information harvesting by automated bots, although these can be an annoyance for normal users however these can be quite usefull when guarding against automated scripts.

6 Conclusions

Online social networks are much vast and much more loose then real life, some people in such networks have hundred of friends in their profiles and may be even thousands through extended profile but hardly any of those are real friends, most of them are complete strangers, and still persol and sometimes sensitive information is freely and publicly available. Based on the information profided in their profiles, users expose themselves to various physical and online risks, these risks are not unique to just one social network, it is common to all of the social networks.

When participating in such social networks, users should be vigilant and should take special precautions before they expose their identities, technology should make it easy for users to protect their privacy, the privacy controls should be well placed an easy to use, so any one can use them.

Social networking is still evolving, new innovations and ideas are coming out rapidly, social networks are now targeting different mediums, mobile media is one such medium, in the future we may see that social networking moves from web sites to handheld devices and technologies such as bluetooth and GPS may play a crucial role in mobile social networking arena, however it remains to be seen how users privacy will be handled when such new ideas are implemented. [9]

References

[1] Captchas: Telling humans and computers apart automatically.

[2] Defenses lacking at social network sites. 2003.

[3] S. Arrison. Its friendster the new tia. 2004.

[4] J. Golbeck and J. Hendler. Inferring binary trust relationships in web-based social networks. *ACM Trans. Inter. Tech.*, 6(4):497–529, 2006.

[5] R. Gross, A. Acquisti, and I. H. John Heinz. Information revelation and privacy in online social networks. In *WPES '05: Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 71–80, New York, NY, USA, 2005. ACM.

[6] K. Jump. A new kind of fame. *The Columbian Missourian*, 2005.

[7] J. M. Kleinberg. Challenges in mining social network data: processes, privacy, and paradoxes. In *KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 4–5, New York, NY, USA, 2007. ACM.

[8] R. Kumar, J. Novak, and A. Tomkins. Structure and evolution of online social networks. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 611–617, New York, NY, USA, 2006. ACM.

[9] C. Lampe, N. Ellison, and C. Steinfield. A face(book) in the crowd: social searching vs. social browsing. In *CSCW '06: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 167–170, New York, NY, USA, 2006. ACM.

[10] A. Lenhart and M. Madden. How teens manage their online identities and personal information in the age of myspace. Technical report, 2007.

[11] A. Leonard. You are who you know. *Salon.com*, 2004.

[12] I. Liccardi, A. Ounnas, R. Pau, E. Massey, P. Kinnunen, S. Lewthwaite, M.-A. Midy, and C. Sarkar. The role of social networks in students' learning experiences. *SIGCSE Bull.*, 39(4):224–237, 2007.

[13] R.Gross. Re-identifying facial images. 2005.

[14] I. Sege. Where everybody knows your name. 2005.