Privacy Concerns in Social Networks and Online Communities

Amirhossein Mohtasebi
Solution Developer
Extol Corporation
amirhossein.mohtasebi@extolcorp.com

Parnian Najafi Borazjani
IT Security Management Specialist
University Technology Malaysia
nbparnian2@siswa.utm.my

Abstract:

This paper studies the awareness of students in Malaysian universities regarding the threats in online social networks, mainly Facebook, and measures their level of vulnerability to those threats, such as identity theft and online harassment. Taking the maturity of social networks into consideration, both in terms of number of users and provided services, the authors intended to find out how concerned university students are about their privacy in online communities, and the extent to which they use the privacy enhancement enablers that are available to protect their personal information.

Introduction

Since their advent, Social Network Sites (SNSs) have attracted different groups' attention. Technical experts, Businessmen, Social Activists, and Researchers try to find out the structure of SNSs, and at the same time take advantage of those sites as normal users occasionally do. In other words, SNSs are an inseparable part of people's life now.

Among all the concerns in the social web, privacy is hotly debated among researchers, legal officials, and social network operators. While providing privacy on the web environment is an enormous ongoing issue for stakeholders, the magnitude of these concerns is more tangible after considering SNS characteristics, which are the number of users, the virtual proximity of them and the huge pile of unprotected information about them, shared with minimum restrictions with known and unknown entities.

Besides the annoying nature of targeted advertisements, there are many privacy issues that threaten personal information, such as stalkers, child and family abusers, employers and even governments (Jacquie Bowser, 2009). The open nature of social networks allows information seekers to find their targets easily by powerful search tools, to impersonate themselves as somebody else or do some sort of identity thefts to bypass privacy options. Finally, they gather and analyse information about their targets, by using profile information, comments, joined groups, and even his/her contacts' preferences.

In this paper, there is an attempt to consider different dimensions of privacy and find out to what extent members of those sites are vulnerable to different social attacks. Additionally, based on a sample set recruited for this survey, we wanted to find a relationship between unique characteristics of respondents (e.g. Age, Gender and Education) and their level of vulnerability and awareness about privacy enhancement enablers available in SNSs. This relationship can show for example whether females are more vulnerable in social attacks, and what information they are willing to share more.

Literature Review

History and Definition of Social Networks

Among researchers, there is a common understanding about social networks. However, based on a definition by Boyd (2007, p. 2) a SNS is a web-based environment that allows users to (1) create a public or semi-public profile within a bounded system, (2) construct a list of users in that system with whom the user maintains connection, and (3) view connection lists (his/her own or friends' lists) and traverse between other connections' profiles.

By taking this definition into consideration, some online applications, such as AIM, (AOL Instant Messaging System) and other messengers, cannot be considered as pure social network but only a part of their structure. However, there are lots of bindings and integrations between available social networks and these systems (O'Hear 2007) (Oiaga 2008). On the other hand, there is not a common way in these environments to satisfy the third element mentioned above, in terms of traversing between profiles.

Many online communities fulfil the aforementioned definition of social networks. However, there are many differences in their geographical distribution, target population and structure. Some of those sites are only successful in some parts of the world. Live Spaces (formerly MSN Spaces), for example, was introduced with the hope of worldwide success in 2004. While its success in its homeland, the United States, was not considered a big hit, it attracts many users in Asian countries such as China and Japan (Boyd 2007, p. 9). Similarly, Orkut, a social network that is operated by Google, was intended to be a worldwide social network, but its current members are mostly from Brazil with 49.66% of users, followed by the United States and India with 20.48% and 18.10% of users respectively (Orkut News 2009).

Additionally, some SNSs have their eye on only a segment of active users on the Internet, for example, LinkedIn and XING target business networks, YouTube targets video sharing and Last.fm activities are dedicated to music sharing. Besides, all of these social networks, there are some high ranked SNSs that target a wide range of users, in terms of demographic and geographic attributes, such as Facebook and MySpace. It is worth noting that an SNS such as Facebook was formerly a dedicated website for college students only (Yadav 2006) and then transformed into a public one (Kornblum 2006). One of the latest hypes was over short messaging networks such as twitter that attracts lots of users as well as celebrities to maintain a very simple and effective means of communication with friends and fans. Despite all the differences in social network sites, based on Gross and Acquisti (2005) the core functionality of SNSs is to provide each user with a profile page, and occasionally the users may give privileges to other users to access their profiles, in order to create new connections (by any means) with new people. However, many users are not necessarily looking for new connections, but trying to maintain their current connection with their friends in the real world (Dwyer, Hiltz & Passerini 2007).

Privacy and Information Revelation in Social Network Sites

Social Network Sites, and mainly Facebook, as our main case study in this paper, attract a huge number of users across the world. Based on a recent report, Facebook has more than 200 million active users (Users who have returned to their profile at least once in a month) (Facebook 2009). This figure transforms SNSs to not just social networking websites, but piles of crude information that can be processed and occasionally abused by other parties such as online hackers, villains, employers, governments, and stalkers. Moreover, the problem deteriorates when normal human behaviour in real life, in terms of creating layers of personal information and sharing them based on the nature of friends, changes in social network communities. A study argues that people are much less concerned about hiding themselves behind those layers and tend to share as much information as possible to their online connections (Gross & Acquisti 2005). This can be the reason why people try to break the taboos and reveal information such as their sexual orientations or their drug usage habits to everyone, or post contents that are not aligned with social norms.

Information revelation to public may have different effects on users' life. While a student or an employee may be fired by college principals or companies' management for undesirable behaviour, in other cases it may lead to worse consequences. The following are minor examples of these threats: a child may be abused by a paedophile who is searching for his victims online, national security may be threatened because of a government employee leaking confidential information

on his or her Facebook page, a celebrity could be blackmailed or at least ridiculed by inadvertently publishing his or her private photographs online.

Social Network Theory and Trust in Social Network Sites

Researchers discuss human behaviour in social environments and by using Social Network Theory. In this theory, they categorise relations and coined the term “weak” and “strong tie” between network members. Based on Kadushin (2004) and Granovetter (1973), “weak ties” and “strong ties” can be defined as below:

“…Our acquaintances (“weak ties”) are less likely to be socially involved with one another than are our close friends (“strong ties”).”

However, based on Gross and Acquisti (2005), the way people behave in online social networks is far different from what social network theory mentioned about “offline”’ social networks. Social network theory defines relationships’ strength in real or “offline” social networks as a spectrum between “weak ties” and “strong ties”, so people share much information to connections with strong ties and are closer to them, while they share less information with people with whom they have “weak ties”. In contrast, Boyd (2004) argues that relationships in social network sites are a binary variable, which is “Friend” or “not Friend”. In our target SNS , Facebook, although users have the ability to create different groups and give them different privileges to access only parts of their profiles, studies show that for the following reasons, it is hard for normal users to use these features.

In Facebook, there is a tool called "Friend List" and users can use that in order to categorise their connections, but, unfortunately, these groups are shaped mostly based on locations and occasions (i.e. School Friends, University Friends, and Work Colleagues) and they just resolve some minor privacy issues. Hence, the different social ties a worker may have with his boss and with his colleague, although both of them are in the same group called “Colleagues”. If the user wants to create restrictions for his boss, he should create a new group. Consider the number of additional groups in one hand, and the amount of effort he should invest on the other hand, to reach to an ideal privacy.

Privacy and Trust in Facebook

Facebook, as the biggest social network site (Kazeniac 2009) and the fourth largest site in the world (Schonfeld 2009), is at the focal point of researchers’ studies about privacy. Large numbers of users distributed geographically and demographically, and the huge amount of information which is posted by them every day, make Facebook a good target for researchers, governments, marketers and hackers. Moreover, based on a study conducted by Dwyer et al (2007), Facebook users have more trust in the environment itself and the users they are contacted by, and as a result they share more information there.

We found three main elements for privacy concerns in SNSs in general and Facebook in particular.

Firstly, in offline social networks, we either share a small amount of public information with a relatively large number of connections or share a large amount of personal information with a small group of connections, mostly friends or legitimate entities such as police enforcement agencies, lawyers, or psychiatrists. While in social networks, by default we share a large amount of information with a large

number of connections. In the Facebook case, while in the first degree of separation you may be connected to your friends or trustees, in the second degree of separation -friends of friends- you may be connected to thousands of totally unknown connections and you inadvertently share a huge amount of personal information with them, such as tagged photos, notes or wall posts. This situation worsens when the information such as a tagged photo of you in a party is not shared by you, but by your first level connection, and by default, it will be accessible to thousands of known and unknown connections. While you have at least some rationale in choosing your known friends and you have some minimum trust in all of them, you have absolutely no control on the behaviour of unknown connections watching your personal pictures, and even using them for any reason. In another perspective, if we define different levels of one's trust in different connections as some circles around him, the farther the distance from unknown connection gets, the lower the trust is.

Secondly, based on research by Gross and Acquisti (2005) that was conducted in universities, most of the students find it difficult to locate privacy options in the Facebook interface and to configure them based on their preferences. One of the reasons is that using those capabilities is a very time consuming, and complicated task as the number of connections increases. Additionally, Strater and Richter (Strater & Richter 2007) argue that most of the college students are not familiar with privacy settings in Facebook.

Finally, privacy settings in Facebook are not as strict as they should be by default. There were many of those loosely set privacy settings incidents in the history of Facebook. In a review by Kirkpatrick (2009), he argues that Facebook set the default privacy settings in such a way that public users have access to wall posts, photos, and videos posted by an individual. Obviously if users are lucky enough to know about this configuration, they can change the option to restrict the viewers to their trusted connections only. One of the real world incidences of this default privacy option was the leakage of personal details and photos of the incoming head of MI5 to the public from his wife's Facebook profile (Greene 2009).

Organization of Research

The main aim of this research is to investigate privacy in online social networks such as Facebook in Malaysian universities. We chose Malaysia for our research, because of some unique characteristics of its society. Firstly, Malaysia is a good example of a developing country that its people are quite new to online social networks. Secondly, contrary to other countries in that level of digital development such as China, and Iran, there is no censorship on social networks by governmental agencies, therefore, more people can have access to SNS. We focused on universities to have a comprehensive sample of participants in terms of gender, education, demographics, and ethnicity.

What makes this research unique among previous works in this area is its focus on a country that is not mature in terms of digital development. This makes it worthy in the way that people are not aware of threats in online communities so they are more vulnerable than people in more developed countries are.

Methods

Recruiting Technique

For recruiting our sample members, we focused on two international universities in the Malaysia region. The selected universities are Multimedia University[i] (MMU) and University Technology Malaysia[ii] (UTM). We tried to have a diversified sample size in terms of education and nationality. The reason for this diversification is the difference between attitudes of different nationalities and ethnicities with different religions and beliefs to sharing their information, some of our respondents had a background in IT, and perhaps they were more aware about privacy issues. We tried to send online questionnaires to our selected sample size and additionally put table stands in university campuses.

Research Design

The research was mainly intended to find out the level of threat to privacy in social network sites. Its aim was to show that while social networks have matured at least in terms of their popularity, many of their users might not be aware of online threats and the privacy options they may use to prevent those breaches.

The survey consisted of three parts. In the initial part, respondents answered some general questions, mainly intended to mine some basic personal information about them, such as education, gender, and nationality. In the second part, the questions are general questions about privacy, not limited to a specific SNS. In the closing part, questions that are specific to the Facebook website is mentioned.

The second phase was about developing a threat model and randomly, without the prior knowledge of the targets, evaluating them against that threat model. The targets were selected randomly from our main sample size in the previous example. In order to respect respondents' privacy in this research, all the information was gathered anonymously. The process of sample recruitment in the second phase was completely random and all the gathered information from successful attacks was only accessible by researchers.

Threat Model

The threat model's objective was to find how many of our targets are vulnerable to our attacks and how easy it is to access private information of a connection. We created a fake Facebook account and tried to access the main information of a targeted contact. We considered an attack a successful one if the targeted contact accepted our fake Facebook user friendship request and it has at least one of the following criteria:

- Access one item of contact information (i.e. Email and IM, Address, Phone Number),
- Access home or work address,
- Access employer name,
- Access personal albums (not profile pictures)

These attacks' intention is to answer the following questions:

- Whether the targets add unknowns persons or not,
- Whether there is a significant effect on the target's decision if she/he has some mutual friends with the fake account.

We classified users based on their gender and sent them friendship requests; some of them got the request from the same gender and the others from the opposite gender. We considered a request as ignored or rejected when it passed the two week period from the invitation.

Analysis

Adding New Connections

About half of our respondents, 48% to be exact, could be the target of some sort of social attacks as they accepted “Friend” requests from unknown people. The other half (52%) mentioned that they never accept “Friend” requests from unknown people, so they seem to be more immune from breach of their privacy. However, the vulnerable part of our sample did not mention that they accept friend requests the first time, but there are some elements that affect their decision. Almost 10% of respondents accept friend requests without any investigation. About 11% of respondents mentioned the gender of the person who sent the request and 9% mentioned the picture of the new contact as a determinable element of their decision. About 21% of the respondents accept a friend request right away when they see a mutual friend in between. About 7% of them have other measures such as location, university, etc to decide about adding or rejecting the request (figure 1). Interestingly, we did not find any significant difference between the attitude of males and females towards gender and the picture of the person who wants to add them as friend. In contrast, we asked our respondents whether it is easy for them to meet new persons in social networks in the real world, and found a very significant relationship between the respondents who found it hard to meet new people and the ones who never accept friendship requests from unknown people.

Continued next page

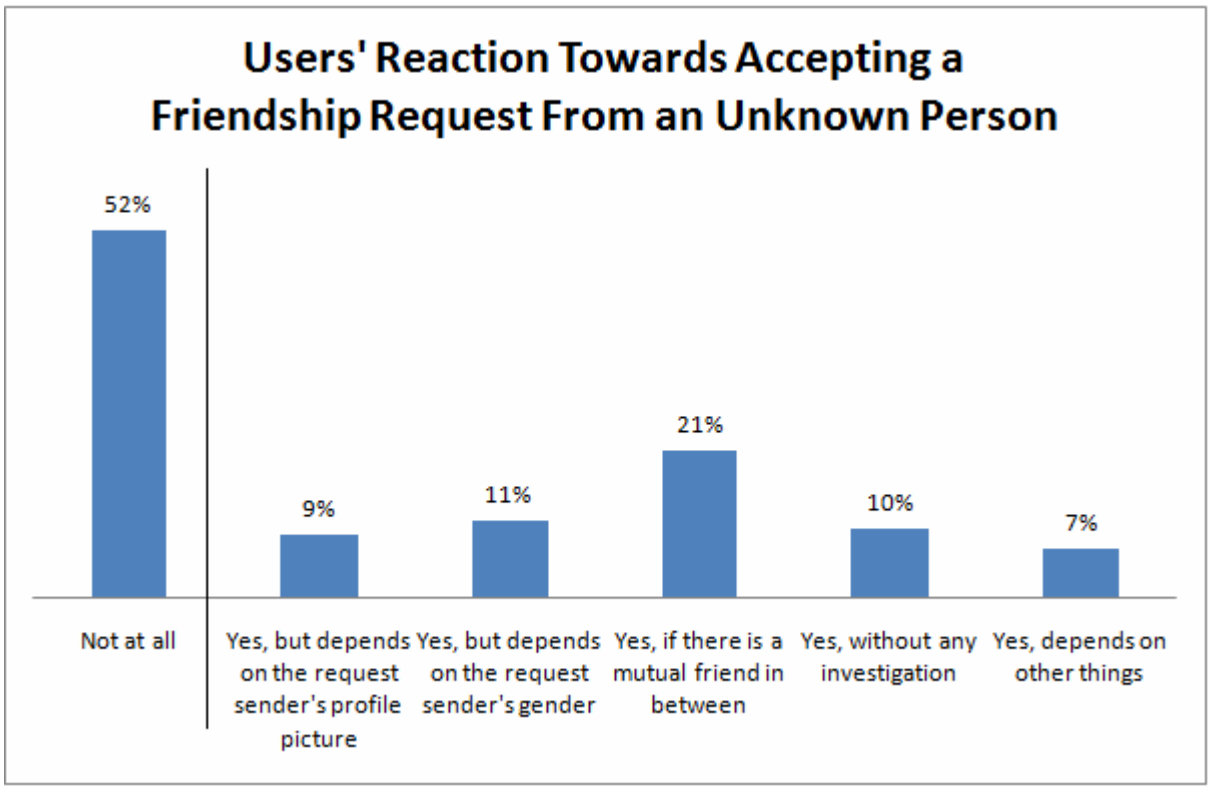

Figure 1 - Users reaction towards accepting a friendship request from an unknown person in social networks

Based on the behaviour of respondents towards accepting or rejecting a new friendship request from an unknown person, we categorised them according to their vulnerability to social engineering attacks. In other words, we considered as vulnerable targets those who accepted the requests without any investigation or being affected by the gender or picture of the attacker, which presumably is fake. Additionally, we assumed a person was a vulnerable target if she/he accepted friendship requests when there was a mutual friend in between. In this way, we found out that there is a significant relationship between the vulnerability of targets and their education level. About 67% of undergraduates are vulnerable, while this figure is about 35% and 17% for postgraduates and PhD students respectively. However, there is no significant relationship between the vulnerability of targets to social attacks and their age.

Privacy and Terms of Services

The respondents were asked whether they read the privacy statements of social networks before accepting them or not. About 14% of respondents had no idea what a privacy statement is and 48% mentioned that they have never read it. Only 5% of respondents care about the privacy statement and always read it before accepting. While reading the whole statement is quite time consuming, 21% of respondents only read parts of the privacy statement. However, privacy statement sections only affected 38% of respondents' decisions on joining a social network.

Utilising Privacy Settings

In all social networks, like Facebook, there are many securities and privacy enhancements available that help users control the visibility of their information and

set what contents can be shown to which group of people. Almost 36% of respondents were not familiar with the privacy dashboard in Facebook at all, or very little, while 28% of them mentioned that they are quite familiar with those options. Forty percent never segmented their friends.

From the respondents who thought they were familiar with Facebook privacy settings more than average, we created a subgroup and tried to find out the degree to which they apply those options in their online social life. About 38% of them had never set the privacy settings as to who can see contents on their profiles' information tab (e.g. their name, birthday, contact name, etc). The figure was more promising for setting what contents should be published on their wall. About 58% of this subgroup used that option for most of their published stories. About 48% of them had never tried to configure privacy settings, concerning whether they would like their personal information to be used in social advertisements. This figure was 55% for the total respondents.

We did not find any significant relationship between being familiar with privacy settings and using those settings.

Vulnerability to Social Attacks

In terms of initiating a social attack, we segregated the attempts into three steps. Firstly, the attacker tried to find the victim. We assumed that the best way of targeting a victim is by using either a search engine, mainly Google, or social network built-in search engine, which is Facebook search in our study. The second step was to convince the target to accept an attacker's friendship request, by using any social engineering techniques. The final step was to access the personal information of the user. However, the third step could be bypassed if the target made his or her information publicly available, so that the attacker could have access to that private information without putting in any effort in step two.

In the first step, 24% of respondents were searchable using the Facebook built-in search. About 5% were searchable using Google search only, although they were not searchable using Facebook search, and 43% of them were searchable by both Google search and Facebook search. The total vulnerable cases in step one was the staggering figure of about 73% (figure 2-a). In order to be more accurate, the respondents were asked whether they use their real identity in social networks or use a pseudonym. While only about 4% never revealed their real identity in social networks, by aggregating this figure with figures mentioned above for search results, we found, to be optimistic, about 70% of cases were totally searchable in Facebook (figure 2-b).

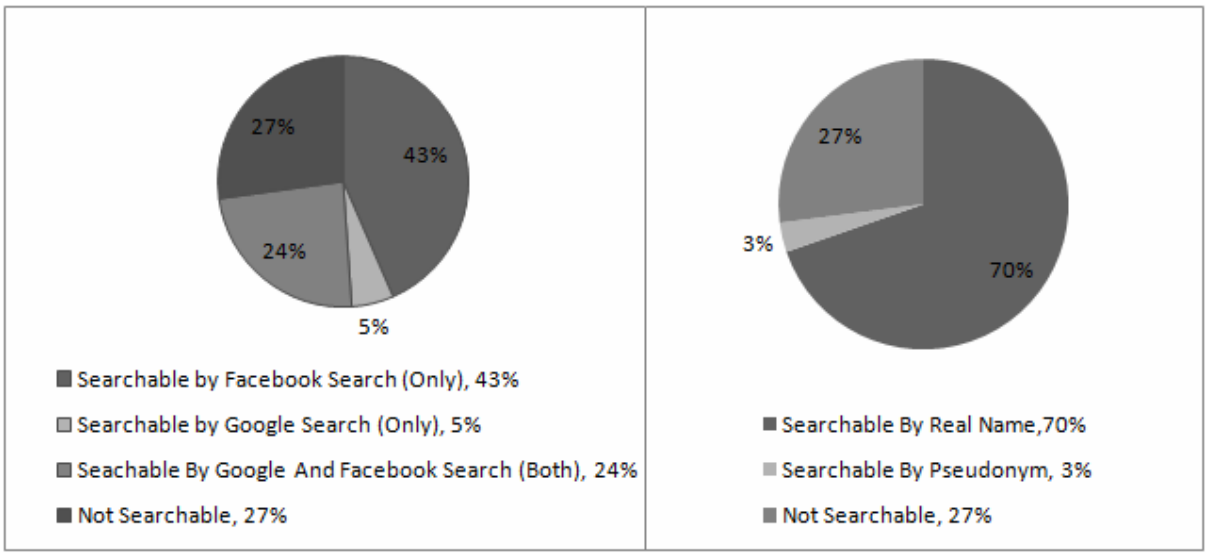

Figure 2- (a) Percentages of users who are searchable (b) Percentages of users searchable by their real identity

In the second step - convincing the target to accept the friendship request - as mentioned above about 48% of our respondents were vulnerable to one or more social engineering attacks. The attacker could put fake profile pictures or gender information to attract the target. Hence, the magnitude of threat, considering that about 21% of respondents accept friendship requests without any investigation when they see a mutual friend. The attacker could try to add one or two friends of the target as friend, to induce the feeling in the target that they were somehow connected to each other in the real world. Additionally, about 84% of our respondents had a positive view about meeting with new people in social networks.

The last step of the attack was to access personal information of the target. Based on our findings, 45% of our respondents did not protect their information by using privacy settings in Facebook at all, so the attacker did not need to make any additional effort to access their information. It is obvious that it does not mean that the other 55% of respondents' information is fully protected from access by attackers. Anecdotally, from the subset of respondents who mentioned they apply privacy settings to protect their information in social networks, about 28% were totally identifiable, and 32% were identifiable by further investigation using information they made publicly available. Finally, about 40% of them could be matched with their identity in real world.

Threat Model Analysis

In order to do threat modelling, an account was created using a pseudonym in Facebook. Afterwards, a subset of randomly selected members from our main sample size was created and friendship requests were sent to them in Facebook. From the friendship requests that were sent, approximately 70% of them were accepted without any question and the other 30% were rejected or are still pending (Figure 3).

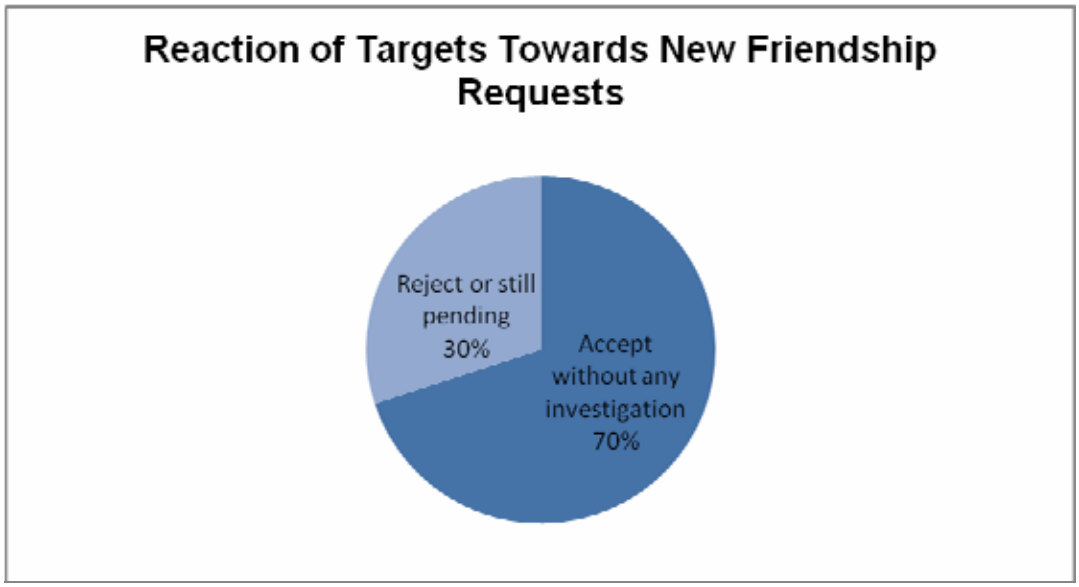

Figure 3 - Reaction of targets towards new friendship requests

In this study, we had a group of 48% females and 52% males. About 79% of targets shared their personal pictures. About 13% put non-personal pictures, and the remaining 8% did not share any picture at all. About 70% of our target group shared their email addresses; that is more frequent compared to sharing their phone numbers (17%) and IM screen name (9%). All these data can be helpful in finding a person via more means and may lead to finding that person even in real life. The other information that may be useful in finding a person in real life includes current address, employer name, and current college or school. Nearly 40% of our targets shared their interest and favourite activities. These kinds of personal information especially can be abused by offenders to seduce young people. In the random group that we chose, at least 9% of targets were under 18 years old; that made them more vulnerable to these attacks. However, in the real world, the number of users under 18 years old is much more than this figure, as this study concentrated on university students.

To put it in a nutshell, in our target group, about 70% of them accepted our friendship request, although they had no idea about our identity. Among the people who accepted our request, about 96% were contactable by using their email address, phone number or IM screen name. However, about 20% of them were contactable in real life using their own or their employee contact information. About 10% of people who accepted the friendship request were contactable directly in real life using their residential address.

Interestingly, within our study period, we got some friendship requests from some unknown people. We had mutual friends with all of them and about 20% of them were from the same gender as our fake Facebook user.

Conclusion

During this research, we tried to concentrate on how successful social attacks can be in retrieving personal information of targets in online communities. First and foremost, we found out that people show different behaviour in online social networks towards creating new connections. Based on previous studies and as endorsed by our research, people seem to be more open in online social networks and are more willing to share information about themselves than in the real world. Secondly, they think they know about potential threats, but they still did not use privacy enablers. We consider it similar to the fact that “Everyone knows smoking kills, but they never regret until they get cancer”. Last but not least, online inhabitants are far more vulnerable to social attacks than we perceived. More than two thirds of our respondents were searchable by their real name; about half of our respondents were vulnerable to at least one of the basic social engineering attack techniques. Consider the rate of successful social attacks when about half of respondents have not used privacy enablers effectively. These figures were endorsed by the experiment we conducted based on our threat modelling technique.

References

Boyd, D. E. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication, 13* (1).

Boyd, D. (2004). Friendster and Publicly Articulated Social Networking. *Conference on Human Factors and Computing Systems (CHI 2004)*.

Dwyer, C., Hiltz, S. R., & Passerini, K. (2007). Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. *Proceedings of the Thirteenth Americas Conference on Information Systems.* Keystone, Colorado.

Facebook. (2009). *Facebook Factsheet.* Retrieved August 1, 2009, from http://www.facebook.com/press/info.php?factsheet

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*.

Greene, R. A. (2009). *Personal details of new UK spy chief on Facebook.* Retrieved August 2, 2009, from http://edition.cnn.com/2009/WORLD/europe/07/05/uk.spy.chief.facebook/

Gross, R., & Acquisti, A. (2005). Information Revelation and Privacy in Online Social (The Facebook Case). *ACM Workshop on Privacy in the Electronic Society (WPES).* Virginia: ACM.

Jacquie Bowser. (2009, March 25). *Big Brother eyes social networks*. Retrieved April 2009, 12, from Revolution: http://www.revolutionmagazine.com/News/MostRead/893804/Big-Brother-eyes-social-networks/

Kadushin, C. (2004). Some Basic Network Concepts and Propositions. In *Introduction to Social Network Theory.*

Kazeniac, A. (2009). *Social Networks: Facebook Takes Over Top Spot, Twitter Climbs.* Retrieved June 12, 2009, from http://blog.compete.com/2009/02/09/facebook-myspace-twitter-social-network/

Kirkpatrick, M. (2009). *The Day Facebook Changed: Messages to Become Public by Default.* Retrieved August 02, 2009, from http://www.nytimes.com/external/readwriteweb/2009/06/24/24readwriteweb-the-day-facebook-changed-messages-to-become-18772.html

Kornblum, J. (2006). *Facebook will soon be available to everyone.* Retrieved February 20, 2009, from http://www.usatoday.com/tech/news/2006-09-11-facebook-everyone_x.htm

O'Hear, S. (2007). *AIM adds further social networking features; borrows from Twitter and Facebook.* Retrieved from http://blogs.zdnet.com/social/?p=325

Oiaga, M. (2008). *Windows Live Messenger 9.0 (2009) Groups.* Retrieved from http://news.softpedia.com/news/Windows-Live-Messenger-9-0-2009-Groups-101043.shtml

Orkut News. (2009). *Orkut News - Demographics.* Retrieved July 2009, from

http://www.orkut.co.in/Main#MembersAll.aspx

Schonfeld, E. (2009, August 4). *Facebook Is Now the Fourth Largest Site In the World.* Retrieved August 5, 2009, from http://www.techcrunch.com/2009/08/04/facebook-is-now-the-fourth-largest-site-in-the-world/

Strater, K., & Richter, H. (2007). Examining Privacy and Disclosure in a Social Networking Community. *Symposium On Usable Privacy and Security (SOUPS).* Pittsburgh.

Yadav, S. (2006). *Facebook The complete Biography.* Retrieved February 2009, 20, from http://mashable.com/2006/08/25/facebook-profile/

---

[i] http://www.mmu.edu.my

[ii] http://www.utm.edu.my