Online Privacy Behavior Laboratory Study- Summary

Andrew Boyd, Todd Kelsey, Joel Silcott, Richard McCarron, Chang Yun, Tariq Wali, Anurag Govil, Jothy Pottukalam, Gopi Keshavan

Seidenberg School of CSIS, Pace University, White Plains, NY 10606, USA

Abstract

This paper continues the study of a relatively new field of online privacy behavior, which has often suffered from a lack of study. This research will study online privacy behaviors from across a wide range of online contexts to identify the source of user initiated disclosure of personal information. We will define online context as sites found on the Internet with specific focus on social and professional interaction, like Facebook and LinkedIn. Our research supports prior published research of similar depth in that online users behave in a manner that suggests that they become part of a group think mentality when part of a social networking group. This leads the researcher to the conclusion that existing models of personal information disclosure are inadequate, and that a new model of online disclosure is required.

Our study contributes to the findings of other researchers involved in similar studies. We hope that the findings in this study can be a significant source in the greater understanding of the differences in disclosure behaviors among users within the social networking world.

Keywords: online privacy, privacy behaviors, online context, social networking sites, safety and security.

1. Introduction

The user of the various social websites may be asked many questions that contain private or personal information. These questions at times can be difficult to answer honestly or, more to the point, completely and accurately. The answers to these questions can be perceived as “half-truths”. A half-truth is a deceptive statement that includes some element of truth. The concepts of half-truths or truths are part of the human psyche that was developed during our upbringing, peer influences, social environment, and/or moral values. The person answering these private or personal questions online may be asked if they answered these questions truthfully 25%, 50%, 75% of the time or not at all. The range of truth may vary according to the topic content of the questions involved. For example, an individual may be more truthful in volunteering personal information on a social network site rather than on a health information or job search websites. Another individual may have many personal reasons for not answering questions truthfully, for example, remembering not being selected for a job or promotion because he/she was too honest in their answers. Also they may not trust the individual or agency that is receiving this information in that it may be used for other means. Other factors that may influence online privacy behaviors are Age, Gender, Education, Ethnicity and Country or Region of the online user. These behavioral aspects are in need of further research in order to serve the global needs of our Internet environment.

The reason for this research is to offer “Proof that our online actions influence the online actions of those whom we interact with on social networking sites.” For example, a friend using a particular social network site gives certain pertinent information to that network, and would your friend’s actions influence you to do the same?

This research study will be a continuation of the “Online Privacy Behavior Laboratory Study Spring 2009” and the Online Privacy Behavior Lab Fall 2009” studies. The previous studies consisted of 10 hypotheses, 7-not supported, 2-supported and 1-not
tested. This continued research will involve evaluating 2-supported hypotheses focusing on Age and Education. The supported hypotheses will be evaluated by conducting online survey questions based on these online actions. The online survey will be created by a powerful and simple program to use called Survey Monkey. This program will provide the features needed to create questions to test our supported hypotheses and to compare the results with previous studies. The results of these survey questions will be compiled and analyzed methodically to draw a conclusion of further support for these hypotheses.

Online Privacy Behavior is an area of study that is fairly new and is in need of further research in understanding user’s behaviors. Understanding social influences (such as our friends and relatives) and how their actions affect others take the same course of action online is an interesting phenomenon and should be studied in greater detail. Further research will contribute to a better knowledge base of understanding online behaviors.

2. Study Components
This study focused on continuing the study of online behavioral attitudes towards privacy on various social networking sites. Previous research defined the social medium as pertaining to four groups; social networking sites, professional networking sites, e-commerce sites, and health information management sites [1]. This further research will continue this study through the use of a distributed survey. The research will include general research gathered about individual attitudes towards behavior during visits of these various social networking sites. To facilitate the distribution of the survey this research employed the use of Survey Monkey, an online survey facilitator. Before the discussion of the findings begin will have a brief overview of the issues concerning privacy in the digital age will be discussed.

3. An Overview of the Privacy Issues of the Digital age
There has been a rapid growth in the amount of users who now belong to what has become known as “social networking sites”. The website Wikipedia defines social networking sites as A social network service focuses on building and reflecting of social networks or social relations among people, e.g., who share interests and/or activities. A social network service essentially consists of a representation of each user (often a profile), his/her social links, and a variety of additional services. Most social network services are web based and provide means for users to interact over the internet, such as e-mail and instant messaging. Although online community services are sometimes considered as a social network service in a broader sense, social network service usually means an individual-centered service whereas online community services are group-centered [1]. Social networking sites encourage users to share personal and professional information but at issue is how these sites collect and use the personal information. The information is then shared among other users, and businesses.

Social network sites include Facebook, MySpace, LinkedIn, WebMD, Twitter and others. Facebook has become one of the largest growing networks with over 300 million members and nearly 50 million status updates per day. Facebook doubles the volume of social networking real-time updates that search providers should incorporate into results [2]. This rapid growth has caused concern for the information that users put on these social sites. There is a great deal of worry about the amount of personal information that could be had with very little effort. There is a perceived privacy threat in relation to placing too much personal information in the hands of large corporations or governmental bodies, allowing a profile to be produced on an individual's behavior on which decisions, detrimental to an individual, may be taken. With this much personal information in one place, there exists the possibility of great risks to privacy. These privacy concerns include stalking, identity theft, legal and administrative repercussions. A user profile on any social networking site could include a person’s name, address, date of birth, age, telephone numbers, family members, religion, marital status, profession and your place of work. With one or more pieces of personal information, much can be done by a predator to harm someone or steal someone’s identity quite easily. It is amazing what and how much personal information that users will place in their profiles or write on these social sites. Personal information is also given during shopping on-line known as e-commerce.

Research has been conducted to determine what types of information can be found on the sites and an attempt made to develop a hypothesis of the various possible categories that a user can fall into based on the extent of information that they are willing to share. There were ten (10) hypothesis that were used and the intention was to confirm if these hypothesis can be supported by research. Do social media privacy attitudes vary by (1) Age (2) education (3) internet experience (4) ethnicity (5) nationality (6) gender (7) individual’s sense of having been a victim (8) group size (9) media exposure to social networking risks (10) online context. Some of these were supported based on the research and some were not supported.

Users have a role to play in the protection of their privacy and what information is revealed on social sites and sites that are used for e-commerce. As a member of one of the social sites, the user should make every
effort available to set controls on what information can be viewed by the general public. It should however start with limiting and restricting what and how much personal information is placed on the sites. Privacy of personal information has been identified as a key ethical issue in the information age [3].

4. World Wide Web Site Popularity: Traffic Data
This section will look at both Facebook’s and LinkedIn’s Page Rank across the vast demographic makeup of Internet users. Previous research has focused on the overall number of existing online social networks however, for this continuing research it is more important to look at the traffic levels of these social networking sites to accurately assess their demographic reach. It is believed that a social networking sites worldwide reach is equal to its popularity. Towards these ends we used Alexa traffic rankings [4], one of several publicly available page ranking sites. Typically these are rankings are commonly used as a general indicator of the amount of traffic a site is receiving.

In the first figure labeled Fig.1 we find that FaceBook has an Alexa Rating of 1 and has approximately thirty percent of all Internet users visiting FaceBook. The second figure labeled Fig.2 is the Alexa Ratings for LinkedIn. Here we see that LinkedIn has an Alexa Rating of 4 and has three percent of all Internet users visiting the site. While there is a large discrepancy between the amount of users in the number one ranking and the number four ranking it does not affect the hypotheses.

5. Selection of Sites
For this research we opted to use the online survey creation tool known as Survey Monkey. Survey Monkey allows for the easy creation of a survey that would give the greatest representation of the behaviors we hope learn about in this paper. For dissemination purposes we choose to use contacts gleaned from various LinkedIn accounts. The decision to use LinkedIn contacts as opposed to those from FaceBook accounts was based upon the fact that LinkedIn contacts can be contacted from outside the social network arena.

6. Facebook
Facebook is one of the foremost social networking websites, with over 300 million active users which represents the highest ever user growth across any social networking site on internet. There are bound to be risks to privacy with this much detailed user information arranged uniformly and aggregated into one place. Users may submit their sensitive data without being aware that it may be shared with advertisers and various other partner vendors. Third parties may build a database from the social networks data to sell. Intruders may steal passwords, or entire databases, from social networking sites. When a new user registers on Facebook he or she is provided with a blank profile template consisting of a number of predetermined response categories. These include “basic” information such as gender and hometown, contact information such as email, mobile phone number, campus room/residence, and “personal” information such as interests, activities, favorite movies/music/books, among others. New users may also upload a profile picture and join up to five networks. The default privacy setting for a new user (“X”) is that the entirety of this information is viewable by anyone in any of X’s networks. X’s profile is also viewable by anyone who is “friends” with X on Facebook. While anyone who is not connected to X via network or friendship cannot view X’s profile, by default they can still locate X using a global search engine as well as view a version of X’s profile consisting of their photograph, name, and network affiliation(s) and even friends and send messages.

7. Creating a Facebook Profile
After a deluge of academic, popular media, and governmental criticism, included a sternly worded directive from the Privacy Commissioner of Canada, Facebook is gradually moving towards an "opt-out"
approach to privacy. This will mean that users’ details are kept private until otherwise stated. Facebook’s historical information philosophy and current default settings are to share almost everything included by users in their Facebook account and partner applications. That means if users have never customized their privacy settings, their Facebook profile is shared with not only all their friends and groups, but also with everyone in every network they have joined [5].

8. Scenarios and approach
For the purpose of our study and to reach out to users Team 10 created accounts on Facebook and LinkedIn. The execution of our plan primarily focused on distributing a survey with direct questions to users. The task included sending a message to all Facebook friends of the team members and requested them to partake in and too solicit others to take the survey. We asked members to evaluate the likelihood of certain online scenarios and privacy behaviors and provide an understanding. While some criticized and other appreciated the medium but the answers were interesting.

A note about Facebook: Facebook members expressed significantly greater trust in both Facebook and its members than that expressed by users of other sites and were more willing to share identifying information.

- Hypothesis: Public sharing of private lives has led to a rethinking of our current conceptions of privacy.

9. LinkedIn
Although a professional social networking site LinkedIn is very much on par with Facebook. Both have are considered to have good privacy. Both Facebook and LinkedIn allow contact with only those who you accept as a discussion partner. LinkedIn is a professional networking site while Facebook brings friends together for more Social reasons. LinkedIn CEO Dan Nye states (when talking about Facebook), that it makes sense to keep one’s professional life separate from the personal. (O’Hear)

10. Creating your profile
LinkedIn allows you to create your profile free of charge. Since LinkedIn is a professional site it is best that you create formal profile that will give contacts the best picture of yourself. Fill out your profile completely. Creating a formal profile doesn’t mean that you can’t say something about yourself. Don’t just stick to your education and jobs or positions you had; be funny, be creative; use simple language. It is important to talk about your professional experience.

One of the nice features of LinkedIn is the “Recommendation” option. This feature allows other LinkedIn users to put in a good word for you, right in your profile. You can even reject these recommendations if you find that you don’t agree them. Avoid putting your e-mail address on your profile as it might make people believe that you are willing to connect with just about anyone. Better let the person interested in you contact the mutual friend in order to get in contact with you.

Using the privacy settings will help the user to avoid spammers or salespeople who may prove only to be a waste of time. These privacy filters are optioned by means of a check list. The checklist gives you the possibility to filter your incoming messages whether you want to be contacted for career opportunities, expertise requests, business deals, job enquiries or just to reconnect with others.

Setting your profile to Full View will allow viewers to access more information about you and search engines will find you easier.

11. Default Settings LinkedIn
By default, your LinkedIn public profile is set to display your full profile information. That means any information provided will be available publicly. More than likely you will notice the very first link returned by Google search on your first and last name is you on LinkedIn and many users had no idea how the other person found them and didn’t appreciate the spam for that matter.

12. LinkedIn privacy behaviors
LinkedIn groups which according to the site are private place on the website where you can communicate by sharing information with your colleague and even with your department head. These groups are automatically populated with LinkedIn users who list their company X as their employee. With these seemingly private groups, employees with less understanding of how third-party communities work may share potentially proprietary information about customers, partners, product roadmaps or even financials. While employees should never share unapproved company information via an external social networking site, LinkedIn’s lack of management tools for an assigned, verified company representative is also disconcerting — especially since these groups seem to be generated largely in part by honor system [4]. Here is a caveat: According to LinkedIn’s help documents for Company Groups: “When someone leaves the company and updates their
position on their LinkedIn profile, they are automatically removed from the company network since the company network is for current employees only. But then “If the person has forgotten to update their LinkedIn profile, you can flag their position to indicate that they no longer work at the company and that they should be removed from the company network. The system will take these flags into account and seek to re-confirm the user’s work email address if they try and access the company network.” The policy is based on assumption that someone has to mark the flag for ex-employee and then LinkedIn takes time to verify the flag, in the meantime employee has access to the private information in the group.

- **Hypothesis:** What constitute privacy behaviors and practices on Professional networking sites?

### 13. Amazon
Privacy in e-commerce has always been a contentious issue as users strive to protect their personal information being misused by others. However, the advent of the Internet and the increasing proliferation of technologies in both the marketplace and workplace have been matched by a heightened awareness amongst users that threats to their privacy exist and must therefore be addressed. Despite the empowering benefits of the Web, consumers are increasingly aware that the technology can also be used by online vendors to collect potentially sensitive information regarding them and that this information can be used without their express permission. For example, online transactions require customers’ to disclose considerably more personal and financial information than they would provide in offline Marketers can use the trail of information that results from such Internet transactions - including information on the customer’s searches, comparisons, product and brand preferences, purchase and post-purchase information - to compose very precise customer profiles in their efforts to continuously learn about changing consumer needs. Websites can collect huge amounts of data from users. Retailers, for example, can track our every click, what we buy, how much we spend, which advertisements we see - even which ones we linger over with our mouse. Sites can easily access your entire web browser history, enabling them to try and guess your gender and other demographic information. They know how much we paid for our house, what magazines we subscribe to, which books we buy and what vacations we take. The company purchases just about every bit of data about us that can be bought, and then sells selections of it to anyone out to target us [6].

- **Hypothesis:** In the Presence of prominent privacy information, users are likely to purchase more if not solicited and phished.

### 14. Hypothesis Development
The IUIPC model draws upon Social Contract theory to present a theoretical framework consisting of multidimensional first and second order elements, as well as a series of demographic covariates. For example, the IUIPC construct states that individual attitudes towards the collection and control of personal information and awareness of information privacy practices constitute a user’s IUIPC profile [8]. This individual IUIPC profile influences trusting beliefs and risk beliefs, which in turn have an impact upon behavioral intent. Demographic co-variants are also related to individual IUIPC profiles [8] [7]. A limited number of demographic co-variants were included in the initial version of IUIPC proposed by Malhorta et al, and this study expands that list to include social media specific co-variants. Our hypotheses posit that specific co-variants are correlated to specific privacy attitudes.

### 15. Methodology and Study Deployment
To ensure we had all the necessary information for this study we created accounts in all the sites mentioned above and started sending invites and friend requests. At the same time we created a survey with about online privacy and behavior and studied how people react to them. We drew qualitative analysis based on the responses to the survey and online literature reviews. Our team developed a survey, which asked users what type of information they make available on social network sites within the contexts previously defined. The survey for our study was distributed via mass email from our individual email accounts, Facebook postings, and professional contacts of individual team members. Users of these social network sites were drawn from friends, family, coworkers, students and classmates. We asked survey participants questions that enabled us determine their privacy attitudes and behaviors relating to our four online contexts. The survey consisted of questions that include users activities on social network sites like Facebook and LinkedIn. In terms of privacy, we also asked users about consumer privacy and security settings. The users responded to each scenario using a seven point Likert scale coded from ‘Strongly Agree’ to ‘Strongly Disagree’.

### 16. Study Design and execution
Privacy disclosure within the four online contexts was broken down into four areas:
To produce our hypotheses, those four areas were mapped against our four online contexts:

- Social Networking Sites
- Professional Networking Sites
- eCommerce Sites
- Personal Health Information Sites

Within each of these contexts, we hypothesized that some behaviors are significantly different from others, and we tested the four behaviors across each context in order to differentiate behaviors among contexts. We have charted out this data and formulated some hypotheses about how likely disclosure will be within these areas. Table 1 values indicate our hypotheses for probability of disclosure. ‘High’ means a hypothesis of high probability of personal information disclosure and ‘Low’ means a hypothesis of low probability of personal information disclosure.

<table>
<thead>
<tr>
<th></th>
<th>Social Networking</th>
<th>Professional Networking</th>
<th>eCommerce</th>
<th>Personal Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Demographic Data</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Lifestyle Data</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Address Data</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Personal Health Information</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

Table 1

Our grid suggests 16 hypotheses (4 x 4). As part of assembling the study design we generated 16 hypotheses defined in the following format.

- Hypothesis 1 (H1): Online users are highly likely to disclose personal demographic data within Social Networking contexts.

Through to:

- Hypothesis 16 (H16): Online users are highly likely to disclose Personal Health Information within Personal Health Information contexts.

17. Survey Results

Our survey was conducted with 149 participants and 56% of them were men and 64% were adults in the ages between 35 and 54. In addition, more than half of survey participants have a Bachelor’s, Master’s or Doctorate degree. Most of our participants were either African Americana or Cacuassian residing in the United States.

Our study revealed some very interesting results about social network use. While most participants answered that they use social network sites to keep in touch with existing friends and family (88%), re-establish contact with old friends (87%), and meet new friends and/or connect with people who share my interests (71%), they also responded that they use these sites for Job Searching (10%) and Business Networking (36%). In addition, our study indicates that 35% of our participants use these sites for dating.
18. Limitations
In view of the scale, scope and further research of this study, there are some limitations that should be kept in mind that apply to all phases of our study.
First, there can be an improvement in the selection of sites and the criteria to assess them. Section 3 of the paper explains the given account of our selection criteria. We have made a considerable effort in identifying all sites that fall into the operational definition of social networking contexts; and it is imperative that we have not considered many other social networking sites because of the time constraints. The user count defined in the sample size is particularly sensitive to cut-off levels because of the high mobility in social networking contexts. Due to the lack of access to many aspects of our study, and challenges in communication in expanding our sample forced us to compromise on the depth of analysis. This study has adopted the point of view that at a later point in time the sample could be expanded more efficiently in breadth than in depth, consider other criteria’s and market dynamics. It might be possible we missed an important evaluation criterion or metadata entry. Our choices were driven by research and analysis and lessons learnt from past studies, trends, and current privacy disclosures in social networking contexts. We also had limitations in attempting to evaluate some more qualitative elements, such as the usability of privacy controls or the readability of privacy policies.

19. Conclusion
Our results support the majority of the hypotheses of our study design. Our in-depth qualitative research study aimed to investigate different attitudes and privacy behaviors in various social networking contexts and their use among 149 users in different geographies. The study included teens, young and adults, users and non-users, and covered predominantly United States followed by Canada. The qualitative nature of this study means that our findings do not necessarily represent of all those who use or do not use social networking sites, the results help us understand and gain better insights around people’s use of social networking contexts, attitudes, behaviors and various privacy disclosures.
Our study reveals most respondents disclosed name, city/town, gender, date of birth, schools and their photos across social networking sites while the same set of users were conscious about sharing personal address, telephone number, event plans, sexual preferences, political views, and religion across social networking contexts.
Our results over time for the three studies reveal significant changes in privacy attitudes and behaviours among respondents over a twelve month period. Specifically, attitudes towards privacy consent release and attitudes towards trust of social networking sites have harden appreciably in the past year.
This initial conclusion from our three studies indicates that temporality plays a significant role in privacy attitudes and behaviours. For example, a user’s willingness to disclose specific information varies over time. The co-variables underlying this temporal shift have not yet been identified, but some possible causes are:
1. Greater user familiarity with privacy and security mechanisms of online interactions
2. Increased user awareness of risk inherent within online interactions
3. Increased user experience with privacy breaches or threats to personal privacy
4. Additional factors which have not yet been identified

That said, it appears that once users have disclosed personal information, they are reluctant to, or do not bother to revise their privacy settings for information disclosure. For example, our graph on the previous page for what information users make available online is almost identical to the response graph for this question from our study twelve months ago.

The implication of this consistency is that—“once the toothpaste is out of the tube”—not only is it difficult to put it back in, most people don’t even bother trying. This element of inertia of privacy disclosure was identified in write-up of our spring 2009 study and characterized as the erosion of privacy attitudes and beliefs; however, additional research indicates that the problem is more complex than a simple downward spiral.

The dynamic within this area of a potential mental model for privacy behavior is that while privacy attitudes may harden over time, the inertia of past bad privacy behavior somewhat mitigates the benefit of more cautious attitudes to privacy. Modeling and testing this dynamic is a fruitful area for future investigative study.

Based upon preliminary analysis, we conclude that two additional factors should be added to existing models of privacy disclosure: Temporality influences privacy attitudes and behavior over time, while the inertia of past behavior mitigates against future responsible behaviours.

References


http://www.pcworld.com/businesscenter/article/190136/yahoo_competes_for_social_web_with_twitter_tab.htm l

[4] Mason, R.O. Four Ethical Issues of The Information Age. MIS Quarterly, 10,1 919860, 4-12

[5] © Alexa Internet, Inc. An Amazon.com Company


[8] Malhotra, N., Kim, S., & Agarwal, J. Internet users’ information privacy concerns (IUIPC): The construct, the