WHAT’S IN YOUR PROFILE? MAPPING FACEBOOK PROFILE DATA TO PERSONAL SECURITY QUESTIONS

Jamie L. Pinchot, Robert Morris University, pinchot@rmu.edu
Karen L. Paullet, Robert Morris University, paullet@rmu.edu

ABSTRACT

This study investigates data privacy concerns regarding the use of Facebook and university students’ awareness of privacy control mechanisms, privacy best practices, and possible consequences of unprotected information-sharing in Facebook. Research found that students share a large quantity and variety of personal data in their Facebook profiles that, if compromised, could be used by an attacker to answer personal security questions setup as fallback authentication methods for other online accounts. The consequences of this type of data breach could include unauthorized bank account access, identity theft, or other crimes. Results suggest that students are generally unaware of best practices for data protection within Facebook, and often misuse available privacy control mechanisms. The researchers conclude by offering a set of best practices Facebook users can follow to help protect data privacy in Facebook.

Keywords: Facebook, social networking, disclosure, privacy, identity theft, personal security questions

INTRODUCTION

Facebook is an online social networking web site that was founded in 2004 at Harvard University [10] and quickly spread to other college campuses [17]. In 2006, the site opened to non-academic users [18] and has since seen an exponential growth in popularity. As of March 2012, Facebook has 901 million active users from around the globe who regularly use the site to stay connected to family and friends, and to share information, photos, interests, and other content that is important to them [10]. The social networking site may owe some of its popularity to its roots in the already existing, geographically-confined, real-world communities that launched it on college campuses. Because of its start as a reflection of these offline communities, users may have developed unique uses for Facebook that were not typical of other online social networks [18, 22].

Facebook has created a paradigm shift in how we communicate with others online. We’ve moved from near-anonymous online interactions to sharing personal details and specific “Likes” using our real names, locations, and even relationship statuses. In contrast, most other online communities and message boards are populated with users who register with aliases such as “NYCgirl” or “NewMom435,” keeping their offline identities separate from those they use online. Facebook is unique in the online space because it requires users to register with their own identity. Even other early social networking sites like MySpace did not have this requirement. Facebook, from inception, made it part of their terms of service that you must use your actual name and email address for your Facebook account, specifically and purposefully disallowing anonymous usage [11, 31]. Studies have corroborated that the data shared in Facebook profiles reflects the actual user, and not an idealized or fictional version of the self that tends to be shared in other online venues [3]. This shift to the use of real-life identities in the online space made Facebook extremely effective as a social networking tool because people could find other people more easily, continuing offline relationships online. But, it also introduced a host of data privacy concerns.

Within a Facebook profile, users can choose to share a variety of data, including employers, schools, hometown, current address, phone number, relationship status, birthdate, names of family members, religion, political views and favorites of all types (music, books, movies, etc.). If accessed, this profile data can be used by attackers for a number of malicious purposes, including stalking or the building of a digital dossier which could lead to unauthorized bank account access or identity theft. There are privacy control mechanisms available within Facebook allowing users to lock their profile down, but the most restrictive setting is to allow access to “Friends only,” and many Facebook users accept hundreds of friends, some of which they may not even know in real life [6].
A major concern, of central importance to this paper, is the quantity and variety of personal data stored in a Facebook profile that could be mined for answers to personal security questions, such as “What is your mother’s maiden name?” or “In what city were you born?” These types of questions are often used as fallback authentication mechanisms for accessing accounts at financial institutions and other web sites [1, 27]. They address personal knowledge in the hope that the answers will be well known to the legitimate user and not easily guessable by an impersonator [19]. An attacker who gains access to answers to personal security questions could use them to retrieve or reset passwords for accounts, thus gaining access to those accounts [1, 27]. The types of information often shared within a Facebook profile are alarmingly well-matched to the types of information asked in personal security questions [27].

RELATED LITERATURE

Data Privacy Concerns in Facebook

Privacy has been a contentious issue for Facebook. In fact, the site’s handling of data privacy issues has been highly criticized in the media and even described as a “trainwreck” [2, 5, 6, 12, 24]. The most egregious offenses noted were data privacy violations that led to incidents involving child predators [20, 28], cyberstalking and cyberbullying [21], financial scams [20, 30] and identity theft [4, 23].

Exposure of content (status updates, photos, etc.) shared within the site has been one of the most sensitive issues to date. When Facebook implemented the NewsFeed feature in 2006, it effectively took all of the information shared between friends one-on-one and brought attention to it by creating a running series of posts showing all activity of all friends in a user’s network. The content that was being shared was already available to all friends in the user’s network, but to see it prior to the NewsFeed feature, users had to go specifically to a person’s page. The NewsFeed called attention to everything by aggregating it all in a central place. This change was so dramatic that it inspired panic amongst users that had to be immediately addressed in a calming blog post by founder Mark Zuckerberg [6, 32]. Boyd [6] likened this experience to singing out loud to a song on the radio in a car full of people only to have the sound abruptly shut off, calling painful attention to your singing voice in the ensuing silence. The more recent Timeline feature introduced in 2011 and rolled out in early 2012 has been received with similar apprehension by many users [26].

Other criticisms have focused on the way Facebook handles privacy setting updates. The site has tended to introduce new privacy settings quietly, with defaults set to the most “open” of options, thus being accused of pushing users to make their data more and more public without their explicit consent [5]. Though the site has continually improved privacy settings and controls for users, critics still take issue with many of the habits employed by the site in relation to privacy concerns, showing a marked lack of trust [9, 25].

Privacy controls in Facebook allow you to restrict access to profile data and content shared within the network. The following options are available for privacy settings, listed from most to least restrictive: Friends Only, Custom, Friends of Friends, and Public. Most of the options are self-explanatory. The custom setting allows designation of groupings of friends to have access to specific types of content [9]. Gundecha, Barbier & Liu [16] note that a single friend in a user’s social network with insufficient privacy settings can place all friends at risk. They contend that a user in a social network is as vulnerable as their least secured friend.

It should be noted that the term “friend” in Facebook can be ambiguous and merely indicates a consensual connection between two users. These connections may not all represent relationships that would normally be described as friendship [6]. In fact, many Facebook users accept acquaintances, friends of friends, and people they do not know in real life as “friends” in Facebook. Boyd [6] argues that while it is not uncommon for Facebook users to have hundreds of friends, those friends are not necessarily close friends, and users are not actually keeping up with the lives of all of them. Further, Boyd [6] claims that Facebook users often treat their friend list as an address book, keeping ties not for the purpose of daily upkeep but rather as a way of getting back in touch with people they’ve met that may prove useful at a later time.
There have been several studies conducted to look at privacy issues related to Facebook and other online social networking sites. Joinson [18] surveyed 241 Facebook users regarding uses, motivations for use, and data privacy issues for the site. The majority of respondents reported that they had changed the default privacy settings on their account. Of those who changed their privacy settings, 25.6% reported making their profile “somewhat more” private, 21% reported making it “much more” private, and 10.9% reported making their profile “as private as possible.” In contrast, 9.2% of respondents reported making their profile “more open” and another 9.2% reported making their profile “as open to others as possible.” The study found that the primary motivation behind making a Facebook profile less private was the desire to meet new people. The average number of friends reported by participants was 124.

Lampe, Ellison & Steinfield [22] studied a dataset of more than 30,000 Facebook profiles to explore the relationship between how much data was shared on a profile and number of friends. The study found that the number of profile fields filled out in Facebook is positively related to the number of friends listed for the user.

Debatin, Lovejoy, Horn & Hughes [8] conducted a survey of 119 college undergraduates to study Facebook information-sharing activities and privacy practices. Nearly 18% of the respondents reported that they had personally been a victim of cyberstalking or harassment, damaging gossip, or theft of data. An overwhelming majority of participants, 91%, claimed that they were familiar with Facebook privacy settings, while only 69% reported that they had actually changed the default privacy settings on their account. Approximately 50% reported that they had set their privacy settings to “Friends Only.” However, 10% of respondents claimed that they accept “anybody” as a friend, and 37% reported that they will accept someone “heard of through others” as a friend. This indicates that almost half of the respondents were willing to befriend strangers. The study found that even though participants claimed to understand Facebook privacy settings, their habits indicate that they have accepted large groups of friends, in some cases including people they have never met personally, and that they share high quantities of detailed personal information. The authors concluded that the gratifications of using Facebook outweighed the perceived threats to privacy for the respondents in their study.

Christofides, Muise & Desmarais [7] surveyed 343 undergraduate Facebook users to explore their habits for information disclosure and control, as well as their personality factors that influence disclosure and control. Participants reported that privacy was important to them, even though results indicated that they disclosed more information on Facebook than they would normally disclose otherwise.

Implications of Sharing Personal Information in Facebook

Arguably, an attacker can hack into a Facebook account more readily than other types of accounts. Social networking sites are not typically considered to need strong authentication protection, so there are often less stringent security requirements in place for them than, for example, online banking accounts. In addition, few social networking sites use SSL (encryption) to prevent interception of passwords. Perhaps of even more concern are the third-party applications, including games, quizzes, news readers, and more that are available within Facebook. Any security breach in a third-party application could potentially expose a large quantity of user profiles or friend lists to attackers. Third-party applications could also be used maliciously to purposefully gain access to user profile data [27].

If a Facebook account is breached, it may also be subject to cross-referencing with other online social networking accounts held by the same person. Gross & Acquisiti [15] note “face re-identification” as a way of linking a Facebook profile with profiles from other online sites for the same user. They claim that many online social network users tend to re-use profile pictures on multiple networks, allowing for easier identification of matching profiles and deepening the level of personal information that can be gathered on one particular user from multiple online social networks.

Mother’s maiden name is perhaps one of the most widely used personal security questions for authentication. It can also be considered one of the least secure, as studies have shown that it can often be discovered from public records.
Social networking sites may render this question even more insecure. Many Facebook users now indicate relationships with family members such as parents, siblings, grandparents, aunts, uncles, and cousins on their Facebook profile. This readily available genealogical data makes it even easier for a data thief to determine an individual’s mother’s maiden name.

Social security numbers are also often used for identification. A social security number (SSN) is not assigned randomly, but rather is based on a number combination indicating the geographical region of birth (often represented by a ZIP code) and an additional string of numbers assigned according to a discernible pattern that can be determined based on date of birth [14, 27, 29]. Therefore, zip code and date of birth, which are often shared on Facebook profiles, could be used by a clever data thief to deduce a social security number [8].

Users can increase the security of personal security questions used for authentication by purposefully answering the questions untruthfully. Using an answer that is untrue renders the answer very difficult to deduce, even for a human attacker with personal knowledge of the user. However, these answers must be memorized and recalled faithfully to be of any service to the user themselves. Rabkin [27] conducted a survey of 46 participants and found that users are typically honest when answering personal security questions. His results found that 38% always gave truthful answers to security questions, 18% seldom lied, 31% sometimes lied, and 13% said they usually lied to purposefully falsify answers (to make them more secure).

PURPOSE OF STUDY

The literature examined above makes it clear that the quantity of personal data stored in Facebook profiles could be instrumental in aiding attacker’s intent upon identity theft, illegal access to bank accounts, or other crimes. If accessed, private profile data could be mined for answers to personal security questions, or used to discover social security numbers. If Facebook users are unaware of best practices for protecting private data in Facebook, their data could be left vulnerable.

This study sought to explore how university students with Facebook accounts feel about data privacy, and to discover their awareness of available privacy control mechanisms, privacy best practices, and consequences of unprotected information-sharing in Facebook.

The following research questions were explored:

RQ1: Are university students sharing information in Facebook that could be used to answer personal security questions?

RQ2: Do university students have an adequate understanding of privacy control mechanisms in Facebook and the possible consequences of sharing unprotected information?

RESEARCH METHODOLOGY

This study examined online information-sharing habits, privacy concerns, and experiences with data privacy violations of college undergraduates with active Facebook accounts at a mid-Atlantic university. A survey was administered to 146 college undergraduates in March and April of 2012 using a convenience sample. Prior to administration of the survey, a pilot test was conducted with 62 college undergraduates to test the survey in February of 2012.

The questionnaire consisted of 23 questions. Survey participants were first asked to indicate age, gender, and Facebook account status. If a participant did not have an active Facebook account, the survey was ended. Those participants with active Facebook accounts indicated basic information regarding Facebook habits, including how often they read and share information on Facebook, types of information and photos shared, and number of friends. Participants also specified the information they added to their profile in Facebook, including employer, occupation,
college/university, major, high school, graduation date, religion, political views, music, books, movies, television shows, games, current city, hometown, gender, sexual orientation, maiden name (for married females), and date of birth. Participants specified whether they were friends with any family members on Facebook and if they had identified those family members on their profile.

In order to better understand participants' practices and concerns related to privacy, they were asked to indicate their privacy settings (friends only, friends of friends/custom, public, or I don’t know), and whether or not they were concerned about sharing too much information on Facebook. They also specified if they had friended people that they do not know in person. Lastly, participants were asked about their experiences with data privacy violations including incidents of identity theft, unauthorized access to bank accounts, cyberstalking, cyberbullying, and child predators.

FINDINGS

Of the 146 college undergraduates who completed the survey, 121 indicated that they had an active Facebook account. The remaining 25 responses were removed from the study. The survey respondents (n = 121) were predominantly male, with 83 male and 38 female participants. Ages ranged from 18-49 years old, with an average age of 25 and median age of 24. For analysis, the participants were categorized into generational age groups. There were 69 participants in the Millennial age group (ages 18-33), making up 57% of the sample. The Generation X age group (ages 34-45) represented 30% of the sample, and the Baby Boomers age group (ages 46-64) was the least represented group in the sample, at 13%. The age breakdown is illustrated in Table 1.

<table>
<thead>
<tr>
<th>Age Groups</th>
<th># of Participants</th>
<th>Percentage of Sample (n=121)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-33 (Millennials)</td>
<td>69</td>
<td>57%</td>
</tr>
<tr>
<td>34-45 (Generation X)</td>
<td>36</td>
<td>30%</td>
</tr>
<tr>
<td>46-64 (Baby Boomers)</td>
<td>16</td>
<td>13%</td>
</tr>
</tbody>
</table>

Types of Information Shared in Facebook Profiles and Possible Links to Personal Security Questions

Research question one asked if university students are sharing information in Facebook that could be used to answer personal security questions. Results showed that students are clearly sharing a large quantity and variety of data in their Facebook profiles, as shown in Figure 1.
Figure 1. Percentage of students who share information in Facebook profiles, by category.

Notably, a majority of students share both their date of birth, 63.6%, and hometown, 76%, in their profiles. Together, these two pieces of data could be used to determine a person’s social security number [8]. On its own, hometown could be used as an educated guess to answer a popular personal security question like “In what city were you born?”

Mother’s maiden name could also be deduced from data shared in Facebook. Within our sample, 21% of married females reported that they share their maiden name on Facebook. This is usually done so that old acquaintances can still find the person without needing to know her married name. If a woman’s maiden name is known, it is easier for a data thief to then find information about her maternal family tree through public records [14]. Facebook users also now have the ability to specify family relationships directly within their profile. Within our sample, 97% of participants indicated that they are friends with members of their families in Facebook and 76% reported that they have identified family members on their profile. If any of the family members identified are from the participant’s maternal family tree, then the participant’s mother’s maiden name could be deduced easily without any need for consulting public records. Many personal security questions have also been known to ask the names of family members such as grandparents, in-laws, or children [27]. A popular personal security question is “What is your father’s middle name?” All of these personal security questions are inherently weak because they could be deduced from public records; they become even more vulnerable if family is identified on Facebook profiles.

Another popular personal security question is “What was your high school mascot?” [27]. The answer to this question could easily be discovered through an online search if a person’s high school was known. In our sample, 85% of respondents indicated that they share the name of their high school in Facebook. Another related personal security question is “In what year did you graduate from high school?” More than a quarter of participants, 38.8%, indicated that they share their year of graduation in Facebook.

Another personal security question is “What is the name of the company of your first job?” [13]. This could easily be discovered by looking at past employers listed on a Facebook profile. In our sample, 45.5% of respondents indicate that they share their employer on Facebook.
Many personal security questions ask about personal preferences such as “What is your favorite music genre?” or “Who is your favorite author?” [13]. Answers to these types of questions may also be readily available within Facebook profiles. In our sample, 51.2% shared their favorite music, 41.3% shared favorite books, 50.4% shared favorite movies, 44.6% shared favorite television shows, and 34.7% shared favorite games.

Even seemingly more difficult personal security questions such as “What was your first pet’s name?” or “What was the name of your childhood best friend?” [13] may be vulnerable based on information shared in Facebook. Though the answers to these questions are not specifically listed as part of a profile, many Facebook users share a variety of photographs with captions and comments attached to them in their accounts. Within our sample, 74% of respondents said that they share photos of their friends, and 43% share photos of their pets. A wealth of information can sometimes be included in photographs, especially photographs that are “tagged” with the names of people or animals in the photo and annotated with captions and comments.

A majority of students, 89%, reported that they have used personal security questions as security mechanisms for other online accounts. Students in our sample are clearly sharing a great deal of information that, if compromised, could be used to answer the personal security questions they have selected and used, which in turn could lead to unauthorized account access or even identity theft.

Students’ Understanding of Privacy Control Mechanisms and Possible Consequences of Unprotected Data

Research question two asked if students have an adequate understanding of privacy control mechanisms in Facebook and the possible consequences of sharing unprotected information. Fifty-five percent of the participants reported that they were concerned about sharing too much information on Facebook, while 45% were not concerned. However, of the 55% who indicated concern, 93% filled out their Facebook profile, 21% shared their Facebook password with a spouse, girlfriend, or boyfriend, and 30% accepted friends that they do not know in real life. A majority, 62%, of these concerned participants reported setting their privacy controls to “Friends Only” which is the most restrictive setting. However, 27% of them have their controls set to the less restrictive “Friends of friends” or “Custom” settings, and 3% have left their profiles completely open to the public for viewing. A small percentage of concerned participants, 8%, indicated that they do not know how their privacy controls are set. The majority of non-privacy-concerned participants, 73%, also indicated having privacy controls set to “Friends Only.” Eighteen percent (18%) reported “Friends of friends” or “Custom” settings, 2% reported the “Public” setting, and 7% indicated that they did not know their privacy settings. Overall, the respondents who indicated that they were concerned about sharing too much information on Facebook actually reported less restrictive privacy settings than those who were not concerned, as shown in Figure 2. All of these factors considered together clearly show that while a number of participants may be concerned with privacy in Facebook, far fewer are aware of available privacy controls and how to best utilize them.
In general, participants did not seem to be aware that limiting the number of friends that they accept within Facebook can directly affect the level of security for their private data. Participants in the study reported an average of 351 Facebook friends, with responses ranging from 3 friends to 1,400 friends (median = 250). Further, 40% reported that they are friends on Facebook with people that they don’t know in real life. This is an alarmingly high percentage of users who have essentially given access to their personal data to strangers. Of the 40% who are friends with people they don’t know, 69% have their privacy controls set to “Friends Only.” While their controls are set to the most restrictive level Facebook allows, they have accepted strangers as friends, thus eliminating the purpose of using this control. In addition, our results show that the 40% of respondents who are friends with people they don’t know share more information in their profiles, in all profile categories, without exception. This indicates that many Facebook users, at least in this sample, may not truly understand the privacy control mechanisms in Facebook or best practices for using them. These results are consistent with the findings of a previous study by Debatin, Lovejoy, Horn & Hughes [8].

Another security concern could be a growing trend showing the willingness of Facebook users to share their account password with friends, family, or their significant other. Within our sample, a surprising 18% of participants indicated that they have shared their Facebook password with their spouse, girlfriend, or boyfriend. Overall, 32% of females shared their password compared to 12% of males. A chi-square test indicated that there is a statistically significant relationship between sharing passwords with significant others and gender (chi-square = 6.684, df = 1, p < .01). Females were more likely to share passwords than males.

Students were also asked about their concerns and experiences regarding consequences of data privacy violations. Of the participants, 60% were concerned about identity theft, and 55% were concerned about unauthorized access to bank accounts. A smaller percentage of the sample had personal experiences in these areas. Six percent of respondents reported being a victim of identity theft and 15% reported experiencing unauthorized access to their bank accounts. Of those who have been victims of unauthorized access to their bank accounts, 17% said that they believed, in retrospect, that information they shared on Facebook was a factor contributing to the incident.
LIMITATIONS

The primary limitation of this study was the small sample size. Because of this, age and gender groups captured within the study may not be representative of Facebook users in general. The researchers encourage future studies to incorporate students from multiple universities or geographic regions to get a better cross-section of participants with active Facebook accounts.

CONCLUSIONS

Use of privacy control mechanisms in Facebook is not enough. Facebook users must understand best practices for the use of these controls, including limiting friends and learning how to share information selectively and responsibly. Attackers can gain access to profile data either via legitimate friend, friend-of-friend, or public access or by hacking individual Facebook accounts or third-party application data. If a profile contains shared data that matches up with personal security questions that have been used at other institutions, it could arm attackers with all of the data necessary to breach sensitive accounts, make illegal financial transactions, and even commit identity theft. Other possible consequences of information-sharing within Facebook include cyberstalking and cyberbullying. Any information shared within the online social network should be carefully considered before being posted.

Facebook users should follow best practices for data privacy which include (1) Limiting friends to a small number of trusted individuals (not strangers), (2) Setting privacy controls to their most restrictive setting (friends only), (3) Selectively sharing only profile data that has not been used to answer personal security questions at other institutions, (4) Carefully considering content (photos, status updates, comments, etc.) before posting it with an understanding of the types of information attackers might gain from it, and (5) Selecting only trusted third-party applications for use within Facebook and ensuring that unused applications have access removed. Armed with awareness about data privacy issues, a savvy social networker can still enjoy many of the benefits of Facebook while taking care to protect sensitive personal data from falling into the wrong hands.

REFERENCES