Social Networks and Context-Aware Spam

Garrett Brown, Travis Howe, Micheal Ihbe, Atul Prakash, and Kevin Borders
University of Michigan
Department of Electrical Engineering and Computer Science
Ann Arbor, MI 48109-2122, USA
{garretto, trhowe, mikeihbe, aprakash, kborders}@umich.edu

ABSTRACT
Social networks are popular for online communities. This paper evaluates the risk of sophisticated context-aware spam that could result from information sharing on social networks and discusses potential mitigation strategies. Unlike normal spam, context-aware spam would likely have a high click-through rate due to exploitation of authentic social connections. Context-aware spam could lead to more insidious attacks that try to install malware or steal passwords. In this paper, we analyzed Facebook, a popular social networking website. Our goal was to determine how many users were vulnerable to context-aware attack email and understand aspects of Facebook's design that make such attacks possible. We also classified different kinds of email attacks based on certain pieces of data such as birthdays, lists of friends, wall posts, and user news feeds. We analyzed Facebook starting from a single university e-mail address to calculate the number of users who would be vulnerable to each type of attack. We found that a hacker could send sophisticated context-aware email to approximately 85% of users. Furthermore, our analysis shows that people with private profiles are almost equally vulnerable to a subset of attacks. Finally, we discuss defense strategies. Some strategies would require users to coordinate their privacy policies with each other. We also suggest design improvements for social networks that may help reduce exposure to context-aware attack email.

ACM Classification Keywords
H.5.3 Information Interfaces and Presentation: Group and Organization Interfaces – Web-Based Interaction.
K.4.1 Computers and Society: Public Policy Issues – Privacy.

General Terms
Human Factors, Measurement, Security

CSCW'08, November 8–12, 2008, San Diego, California, USA.

Author Keywords
Facebook, email, privacy, social networks, spam

INTRODUCTION
Social networks, such as Facebook, MySpace, LinkedIn, Friendster [3], and Tickle have millions of members who use them for both social and business networking. In this paper, we analyze the extent to which the current usage of such systems also poses significant risks for its users to cleverly targeted, context-aware email spam. We then use the analysis to suggest risk mitigation strategies for both users of social networks as well their designers.

Email is one of the most common delivery mechanisms for network-based attacks such as phishing [4,5], computer viruses, Trojan horses, bots, worms, browser exploits, key loggers, spyware, adware, and others. MessageLabs, which scans millions of emails per day as part of its managed email security service, previously reported that on average every one in 12 emails contains malware [8]. Some studies indicate that users are increasingly comfortable with email [7], including spam-filtering functions. Thus, the success of delivered attacks is dependent almost entirely upon the click-through rate of the email; if the target does not click on the malicious link presented in the email, then the attack usually fails.

To improve click-through rates, many techniques of varying complexity exist that mask the malicious nature of the email: hiding the destination of hyperlinks, falsifying header information, and creative use of images are a few of these methods [21, 14]. Simple social engineering techniques can also help disguise malicious emails, such as by mimicking notification emails from popular financial websites (e.g. bankofamerica.com, chase.com), or by mimicking notification emails from popular commercial websites like amazon.com and ebay.com. However, these social engineering techniques are relatively crude; they blindly send messages claiming to be from large websites to people who may or may not have accounts on those sites.

The attacks we analyze in this paper involve the use of email messages that take advantage of some shared context among friends on a social network such as birthday celebrations, living in the same home town, or participating in a common event. This shared context dramatically increase email authenticity, more easily bypassing spam

filters and increasing the click-through rate for spam that contains advertisements, installs malicious software, or solicits sensitive personal information. This paper examines users' privacy policies within a large University Facebook network and looks at the extent of vulnerability to context-aware email attacks.

We analyzed over 7000 profiles on the University of Michigan Facebook network. Many social networks, including Facebook, do not, by default, make email addresses visible on public profiles. However, we found that there was sufficient information available on public profiles to look up emails in another public database for approximately 90% of the Facebook users in our dataset. We also found that almost two-thirds of users have privacy policies for their profiles that permit access to non-friends. On most publicly-accessible profiles, contextual information exists that can be used to generate context-aware attack email. Overall data from out test network indicates that a spammer could target 85% of users who have visible profiles with context-aware attack email. Perhaps somewhat surprisingly, our findings also indicate that almost an equal percentage of users with *closed (private) profiles* may be vulnerable to attacks.

To help mitigate the risk of context-aware spam, our findings indicate that privacy controls must be viewed from a community perspective – users must collaborate to stop attacks. We also suggest possible changes at the interface of social networking systems that would make it harder for an attacker to reconstruct the relationships in social networks that are a basis for these attacks.

The rest of the paper is organized as follows. We first discuss our methodology as well as the threat model. Next, we classify various types of context-aware email attacks. Then, we present a vulnerability analysis of users in a test university network to different types of context-aware email attacks. Next, we discuss potential defense strategies for both users and designers of social networks. Finally, we conclude.

RELATED WORK

Security companies have pointed out that social networks are likely to become a target for attackers [20]. Traditional attacks have been to corrupt data at sites so as to induce users to click on malicious links. This paper examines the possibility of email-based attacks on users of social networking sites.

Earlier studies have shown that email attacks using contextual information have a higher click-through rate than normal email attacks. For example, Jagatic *et al.* harvested publicly available information from Facebook and used it for email-based attacks [13]. By simply pretending to be a friend of the sender, a phishing attack email had a 72% click-through rate versus 15% for email from non-friends.

There is also evidence that email-based attacks using Outlook address books as an attack vector are very effective. For example, the study in [19] analyzed the graph of relationships implied by address books at a university system and the potential for email worms to spread via the address books. One of the findings in that study was that random vaccination of computers using antivirus software is likely to be ineffective because the outbreak drops very slowly as vertices corresponding to the users are removed from the connectivity graph.

This paper builds on the above prior work. It gives insights into aspects of the design of social network systems that contribute significantly to the potential of context-aware spam. Furthermore, based on the analysis, it suggests design strategies for social networks that would help to mitigate the risks of attacks outlined in this paper.

METHODOLOGY

In our work, we used a five-step approach to characterizing the risk of context-aware spam and establishing counter-measures:

1. Sample data from one popular group on a social networking site.
2. Study the available attributes on public profiles within that social network and determine how they may be exploited to generate context-aware spam
3. Create sample attack emails to provide evidence that it would be trivial to automate the generation of context-aware spam using a database of profile information.
4. Analyze the social network to determine the fraction of users that could be targeted by various attacks, along with characteristics in their profiles that contribute to their vulnerability.
5. Based on the anatomy of attacks, recommend defense strategies to mitigate the risk of context-aware spam.

As in any vulnerability analysis work, we are aware that the results of this paper could incite potential spammers, especially if defenses are not deployed. However, we believe that spammers may already aware of how contextual information can improve click-through rate. It is important to quantify the risks and suggest possible solutions before this type of attack becomes widespread.

Choosing a Social Network

We considered several online networks as potential data sources for our study, including Facebook, MySpace, and LinkedIn. We decided to focus on Facebook because it is popular among university students (over 85% of all college students use it [1]). Facebook profiles are usually well-maintained and also contain large amounts of personal information. Furthermore, a recent study shows that on a scale from 1 to 5, respondents indicated with high confidence (mean=4.16) that their Facebook profiles described them accurately [16]. Although a survey showed

that users trusted Facebook more than MySpace by a very slight margin, the amount of correct information they disclosed on Facebook versus MySpace was significant (100% versus 67% revealing their real name in their profile) [6]. Furthermore, 94% of users disclosed their email addresses versus 40% at MySpace [6].

Facebook Privacy Policy

Some of the important attributes that Facebook profiles contain are: contact information, place of residence, educational background, gender, interests, names of friends, birthday, wall posts (messages posted by other people on a user's profile), and news feeds (list of recent activity by the profile owner, such as event RSVPs, etc.).

Facebook provides simple privacy settings at multiple granularities to control what information other users can see in a profile. The default policy is for almost the entire user's profile to be available to both friends and to people in the user's network (i.e. the entire university, city, or workplace). Users can join multiple networks, including non-university networks. In our study, we only examined the possibility of attacks within a single network.

Closed and Open Profiles

Facebook allows users to completely restrict non-friends from seeing their profile. From the perspective of a stranger (e.g., an attacker), such profiles are unlikely to be accessible. We therefore define such profiles as being *closed*. All profiles that are not closed, meaning a stranger in the network could see the profile, we define as *open*.

Fine-Grained Access Controls

Facebook allows users to have fine-grained privacy controls on their profile attributes. For example, users can restrict certain sections of their profile only to their friends (e.g. only display contact info to friends), while still allowing anybody to see their profile page. However, it was noted in a study by Harvey Jones that even when profiles were completely closed, individual profile attributes can still be found via an advanced search by anyone in the same network, in essence allowing database reverse engineering [15]. Even without using advanced search, Facebook provides significant navigation capabilities. For example, clicking on a birthday in a user's profile allows one to find users on the same network who share that birthday.

Overall, Facebook's privacy model did not significantly hinder our data-gathering tools. While Facebook does allow users to secure their profiles to a laudable extent, the default behavior makes many attributes in the profiles open to all other users in the network. Many students and young people do not take the time to change their default settings, believing that the information they post on Facebook (or MySpace) is private or should be considered private [10].

Test Social Network

For our study, we chose to focus on users in the University of Michigan Facebook network. This network potentially contains anyone who registers using an email address at the `umich.edu` domain. This includes current students, staff, and faculty. As at many universities, this also includes alumni of the university, who are allowed to retain their umich.edu address.

Amenability to Automated Analysis

When looking at the possibility of automatically analyzing profiles, as one would expect an attacker to do, we determined that the HTML for Facebook profiles is fairly consistent and could be parsed automatically, although this was complicated somewhat by third-party applications.

We did anticipate some difficulties in extracting email addresses in an automated fashion from Facebook profiles. Facebook allows users to keep email address as private. Even when the email address is public, Facebook profiles display the email address of the user as a PNG image instead of text. We were unsuccessful in using optical character recognition software to accurately interpret the email addresses (our attempts were admittedly crude in that we used standard OCR software available to us). Ultimately, however, we found that for the network we selected, the first and last name was enough to link 90% of the profiles to an email address with a university directory, which we deemed as sufficient coverage.

Model for the Capability of the Attacker

For the purpose of our study, we assumed that an attacker could find a way to join a target network on Facebook. However, we do not require the attacker to have any friends in the network. For an attacker to gain access to Facebook as a user in the University of Michigan network, which was our test network, the only barrier would be in acquiring access to *one* email address that ends in umich.edu. This would be easy for attackers to do if they were current or former students at the University. Even if that is not the case, an attacker would only need to acquire access to an existing email account, which is not hard to do by compromising a machine on a university network. That email account can be used to either set up a new profile at Facebook or, if a profile already exists, to reset the password and get access to the existing profile. Another alternative is to create a new user account after compromising a machine with a hostname in the same domain (e.g., attacker@victim.umich.edu). Facebook only requires the hostname to end in umich.edu. We suspect that many other educational networks at Facebook, if not all, would be susceptible to a similar line of attack.

We also assume that the attacker can find a way to map most of the full names at Facebook profiles to email addresses. Like many universities, the University of Michigan provides a publicly accessible web-based directory for looking up contact information of students and employees at the university. The directory server does have standard defenses to prevent automated programs from collecting email addresses. For example, it limits the

number of results returned in the case of approximate matches. However, such defenses did not prevent exact lookups on names of users in the University of Michigan network at Facebook. We succeeded in matching 90% of the profiles at Facebook in our test dataset to unique email addresses. The directory does allow users to specify more strict privacy settings, but the default, which most people use, is for the full profile to be visible to everyone.

For Facebook networks where the corresponding university directories have good privacy settings, the university does not have a directory, or for networks not affiliated with an e-mail domain, one could guess e-mail addresses based on a naming scheme. For example, [firstname].[lastname]@ webmail.com or [firstinitial][lastname]@univname.edu. Furthermore, one could extract many e-mail addresses directly from Facebook using advanced text-recognition software that is designed to defeat CAPTCHAs. Mailing list vendors who sell huge lists of email addresses along with full names can be found easily doing by doing a web search. These lists are available to spammers and may also be an effective way to obtain email addresses given a person's first and last name.

Seeding Our Initial Search

One problem that we had to solve was how to start our search. Social network sites are designed to help a single user find a relatively small number of other users with something in common. An attacker needs to find a large number of users with which he or she has no prior connection.

We had initially hoped to use the University of Michigan online directory to bootstrap our search since it contains every student at the university. However, we were unable to do so because the directory severely limits the number of search results to make email address harvesting more difficult. If the size of search results exceeds a threshold, the directory returns no results.

The method that we ended up using to bootstrap our search was Facebook's "Browse My Networks" feature, which returns 10 random users from within a specified network on the first page. By visiting this page hundreds of times, we were able to obtain profiles for a large number of unique users, which as an added bonus should be a random sample of the population (assuming Facebook is selecting the search results randomly). At the time of our study, the University of Michigan network had 70,382 users (Facebook reports this number when you join a network). We collected profiles for 7,919 of those users in about 20 hours, which we considered a sufficiently large group to provide us a representative sample for the purposes of our study. (We chose to do this collection manually and limited the number of profiles collected so as not to violate Facebook's usage policies on automated data collection.) We subsequently discovered that Facebook actually returns about 200 random users on each query; only 10 are displayed on the first page. Thus, it is very likely that an attacker, who would not care about Facebook's usage policies, could easily use an automated script and collect profiles on almost all the 70,382 users in a fairly short time frame. Assuming that Facebook is returning 200 random users on each "Browse my Network" operation, and given a network population of N, after m browsing operations the probability of a particular user among N *not* being selected is $((N-1)/N)^{(200*m)}$. For N being 70,382 users and m being 2000, this value is 0.0034, indicating that most users could be discovered using about 2000 lookups that return 200 random users each (400,000 total results).

CLASSIFICATION OF CONTEXT-AWARE EMAIL ATTACKS

We examined information available on profiles and selected attributes that could be used as a basis for context-aware attacks. We identified three kinds of attacks:

1. *Relationship-based attacks:* These attacks only use friend-to-friend relationship information. No other attributes from users' profiles are required to carry out an attack.
2. *Unshared-attribute attacks:* These attacks use friend-to-friend relationships, along with an attribute from only one of the parties in the relationship. An example is the use of a birthday attribute from one of the parties for devising an attack.
3. *Shared-attribute attacks*: These attacks use friend-to-friend relationships, along with an attribute that is visible at both parties in the relationship. Usually, if the attributes share a value (e.g., common hometown), that could help devise an attack.

Intuitively, we expect that unshared-attribute and shared-attribute attacks will have a higher click-through rate as they provide more authentic context to the recipient. For social networks, relationship-based attacks are a special case of shared-attribute attacks, since the sender and the receiver do share some implied attributes: they are both users of the social networking site and they belong to the same network. The attacks we consider in this paper use that strategy to convert relationship-based attacks to shared-attribute attacks.

Next, we give examples of email templates that demonstrate each of the above attacks and illustrate that they could be made to look very similar to normal messages that the users receive on a regular basis.

Email Template for Relationship-Based Attacks

If we only have friend information and no further social context from a user's profile, then the standard issue email from Facebook that can be seen in Figure 1 could serve as an easy, believable medium to deliver an attack.

[SENDERNAME] has written something on your wall.

To see what [SENDERNAME] wrote, follow the link below:
http://univname.facebook.com/profile.php?id=123
(Link to attacker-controlled site)

Thanks,
The Facebook Team

___
Want to control which emails you receive from Facebook?
http://www.facebook.com/editaccount.php?notifications
(Link to attacker-controlled site)

Figure 1: Sample Facebook Notification Attack.

In this case, the relationship-based information is mapped to a shared-attribute attack, where the attribute is the fact that two friends belong to Facebook. Note the above message in Figure 1 is similar to a notification from Facebook. The only difference is that the link to the user's profile is malicious; it points to a page that is controlled by the attacker and that will possibly inject malware into user's browser, display ads, or induce user to leak personal data.

In a similar way, the attacker could try to leverage knowledge of the common network between two friends, even if profiles provide no other helpful information. For friends on a University network, the attacker could send an email that claims to be a link to information related to the University.

Unshared-Attribute Attacks

If users post their birthdays on their profiles, which 87% of users in our dataset do, then a *birthday greeting attack* like the electronic birthday card seen in Figure 2 is possible. This attack obviously has some temporal limitations; it must be sent near the victim's posted birthday. However, this also adds to its believability. The template for this message was taken from a real email from BirthdayCards.com. We suspect that most users would find it hard to resist the temptation of looking at an e-card from a friend. Again, the attack here is that links in the email would take the user to a malicious site, possibly designed to inject malware into the user's browser or conduct further social engineering attacks.

Besides birthdays, other attributes could be used as well. For example, Facebook has recently started posting users' recent purchases at selected websites to their profiles. If this information is accessible to an attacker, then it could be used to craft an attack email referencing the recent purchase as the context.

Interestingly, public visibility of some attributes on a profile makes not only the profile's owner vulnerable, but also the owner's friends. This vulnerability arises even if the owner's friends make their profiles private. The date of birth attribute is a good example of that. The attacker could conduct a *birthday invitation attack* by crafting an email

Hi [FIRSTNAME],
[SENDERNAME] ([SENDEREMAIL]) has sent you an online greeting card from BirthdayCards.com!

To pickup your card, please click on the following link:
http://www.birthdaycards.com/pickup?ID= A222-FHRE
(Link to attacker-controlled site)

If you are unable to click on the link above, please try cutting and pasting the URL into the address bar of your web browser. You may also go to our website at: http://www.birthdaycards.com **(Link to attacker-controlled site)** and choose the "Pickup" option at the top of the page.
Your Pickup ID is: A222-FHRE

BirthdayCards.com - High Quality Greetings for All Occasions.
If you have any other questions or problems, please visit our support page at:
http://www.birthdaycards.com/support.momd

Figure 2: Sample Birthday Attack Template.

inviting the owner's friends to the owner's birthday party. The invitation would a link to a page that claims to be the map of its location. As long as the attacker can identify the emails of the owner's friends from their names, access to their profiles is not necessary. On Facebook, if a person makes their profile page private, only the person's profile is hidden; the person's name continues to be visible on public profiles of his/her friends, including on wall posts and friend lists.

Shared-Attribute Attacks

Shared-attribute attacks utilize a specific shared social context. This can be anything from an event that two friends attended, to a shared "network" like a workplace or hometown. It is even possible to get images of people who are conveniently "tagged" for easy scraping. Figure 3 shows a potential attack template, based on an email from a photo gallery site. A friend can share their online photo album and send this email to the victim. Note that the fact that two friends participated in an event can be gleaned from their wall posts. Facebook also allows users to post photos that are linked to events. All this information can provide a rich shared context for generating authentic-looking emails.

Not shown are two other shared context attacks that we analyzed. If friends share a home town, it would be possible to send them a cordial email including a link to a news article from home. Friends from a similar network, a workplace for instance, can be similarly targeted with a news article about the network in question.

DATA ANALYSIS OF THE CANDIDATE SOCIAL NETWORK

Recall that we collected 7,919 random profiles, as returned by Facebook's "Browse My Network" feature from the University network. For the 7,919 users in our sample

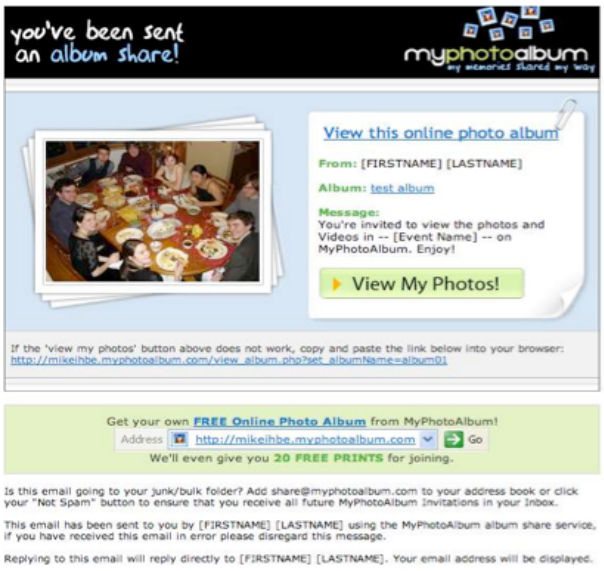

Figure 3: Shared context attack template – a photo from a shared event.

dataset, 5,399 users, or 68%, had open profiles (those accessible to everyone in the network). We thus use 68% as the probability estimate for a profile to be open. Given the data set size of 7,919 profiles, and assuming that profiles returned were randomly selected by Facebook, the 95% confidence interval around this value for an estimate of percentage of open profiles in the entire network is 68.18±1.03%, which we consider to be reasonably precise.

A study by Harvey Jones found that even when profiles were completely closed, individual profile attributes could still be discovered through an advanced search by anyone in the same network, allowing database reverse engineering [15]. The study in [16] also showed that profile structure may be related to friendship links. Thus, 68% could be a conservative estimate for a determined attacker. A higher percentage of open profiles would simply make the attacks even more effective than the values estimated in this paper.

We estimated the usage of the fine-grained privacy controls by measuring the number of attribute fields (out of approximately 35 fields parsed) that were actually present in each profile. As can be seen in Figure 4, a large fraction of users provide access to many fields in their profiles, indicating that they disclose more than just contact information. Keep in mind that not all fields are necessarily present even in open profiles, because setting them is optional. Overall, our results strongly suggest that many users allow non-friends to see attributes other than just contact information. Our findings are largely consistent with an earlier study that also found that users reveal astonishingly large amounts of information [9] on Facebook.

Frequency of Attack Attributes

We also gathered statistics on selected attributes that we considered potentially useful for attackers. Out of the 5,399

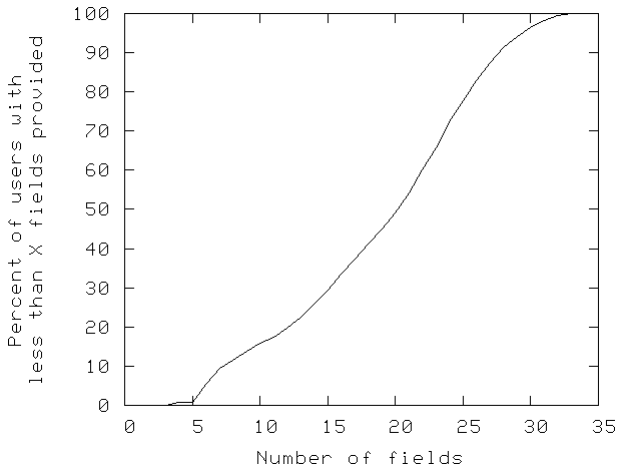

Figure 4: Data available in open profiles.

users for whom we had full profiles, 2,787 (52%) provided the gender in which they are interested for relationships. A previous study has shown that users are up to 15% more likely to click on links in emails sent by members of the opposite sex [13]. We theorize that attackers may be able to improve this figure by using senders in whom the recipient is likely to have a romantic interest.

Of the 5,399 users in our dataset who had open profiles, 4,686 (87%) provided their date of birth. This is probably due to an artifact of Facebook's registration procedures and disclosure policies. When a new user registers at Facebook, date of birth is requested, which Facebook may be using to verify that the user is over 13. The date of birth, if provided at registration, is automatically included in a user's profile. As we showed in our sample attacks, this design artifact could be particularly devastating for attacks around a user's birthday, when friends are more likely to initiate new email communications.

Among the 5,399 users in our dataset with open profiles, we were able to match 4,870 (90%) of them to a unique University of Michigan online directory account, which gives us a valid email address for those users. Recall that we were forced to match based on full name due to Facebook's encoding of email addresses as images. Because names are not completely unique, we were not able to make perfect matches for all of the users with open profiles, but 90% is a very favorable success rate for a potential attacker. For most of the users in the remaining 10%, we got multiple matches. Potentially, an attacker could still achieve success by targeting each matched user in some of the attacks. In our study, we conservatively assumed that the matching success rate is simply 90%, though it could be lower in practice for non-university networks.

We computed the distribution of the number of friends people have in the test network. Users who have no friends are not vulnerable to attacks discussed in this paper. Users with more friends are likely to be more susceptible to

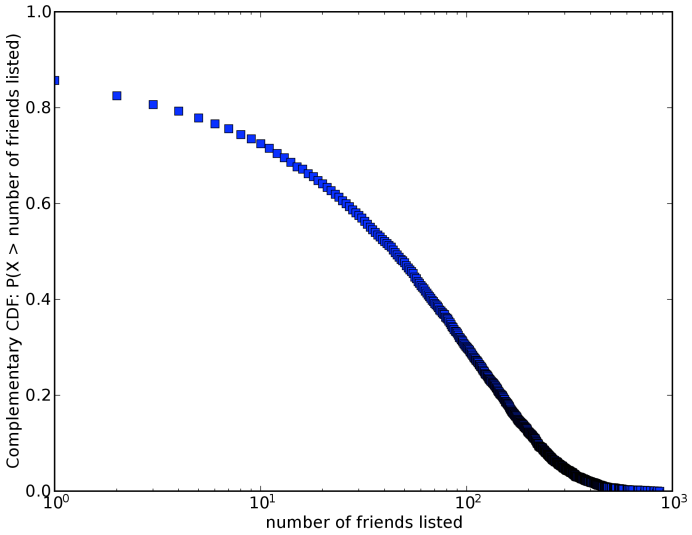

Figure 5: Distribution of the number of friends for users. Approximately 86% of users (with open profiles) had at least one friend listed.

attacks because there will be higher odds of finding friends who disclose their birthday or share events with each other. Figure 5 shows the complementary cumulative probability distribution of friends for users with open profiles in our test network. As expected, it is a heavy-tailed distribution, with some users having a large number of friends (some users had over 800 friends), while around 14% of the users had no visible friends. The friend numbers could be conservative since some users could have made their friends attribute private. A determined attacker could possibly discover hidden friends by examining the names of users in wall posts. In that case, more users could be susceptible to attacks than indicated in our study.

We further analyzed the data shown in Figure 5 to determine if it satisfied the power law property as found in the analysis some networks [2]. It was found to not satisfy power-law properties, which is not surprising. Several networks in previous studies have been found to differ in from a power law curve [11,12], such as exhibiting a faster decay near the tail. The average number of friends per open profile in our sample dataset was 82.

Having obtained some estimates of distributions of open profiles, visibility of attributes, and distribution of friendships, we computed the following probabilities:

- P(open) = Probability that a user's profile is open.
- P(closed) = 1 – P(open)
- P(email) = Probability that a user's email can be deduced.
- P(birthday) = Probability that a user's birthday is visible, given that the profile is open.
- PF(i) = Probability that the number of a friends of a user in a network is *i*.

From our sample dataset for the target social network, estimates for the values are as follows:

P(open) = 0.68; P(closed) = 0.32; P(email) = 0.9;
P(birthday) = 0.87; PF(i) as estimated by Figure 5.

Given these estimates, we next estimated the probability of a random user in our test network being susceptible to a relationship-based attack, birthday-greeting attack, and birthday-invitation attack, based on the following attack models (analysis details omitted for brevity):

- A user can be targeted for a relationship-based attack if the email addresses of both the user and a user's friend are available. In addition, one end of the relationship needs to have an open profile so that friendship can be identified.
- A user can be targeted by a birthday greeting attack if a user has an open profile with an exposed birthday and the attacker has managed to determine at least one of user's friends' email addresses, which he or she will use to construct a forged email. If a profile is closed, the attacker will not have access to user's birthday and a greeting attack is considered not possible.
- A user can become victim of a birthday invitation attack if one of the user's friends makes his/her birthday visible. Assuming the attacker has downloaded all available profiles on the network, the attacker can construct the friend-to-friend relationships even if one end of the relationship has a private profile. Under that assumption, a birthday invitation attack will work equally well on users with open profiles as well as closed profiles.

Table 1 shows the results for our analysis. An interesting finding is that a significant fraction of users with closed

	% Open	% Closed	% All
Relationship-only attacks	85%	84%	85%
Birthday greeting	74%	0	50%
Birthday invitation	84%	84%	84%

Table 1: The percentage of users who are vulnerable to three types of context-aware email attacks based on profile openness. "% All" represents the percentage of all users (with closed or open profiles).

profiles are also vulnerable to attacks because one of their friends has an open profile. Many users in our sample dataset appear to have a large number of friends in their network, making it more likely that an attacker will be able to find at least one friend of a user with an open profile. The lesson here is that users must not limit access to their profiles but also choose friends who are equally careful about their privacy settings.

To analyze shared-attribute attacks, we examined 905 of the users in our dataset who had both open profiles and had friends listed on their profile. We then looked for any shared events among the users and their friends. We found that approximately 228 (24.1%) of them would be vulnerable to a *shared-hometown* attack and 32 (3.4%) of them would be vulnerable to a *shared-events* attack. Even though the percentage of users subject to a shared-event attack appears to be low, a user who was previously not subject to the attack could become vulnerable later on as new shared events are posted. For this part of the evaluation, we only parsed profiles of friends for hometown and event information who were listed on the main page of a user's profile or appeared in the user's wall posts. Thus, these numbers could be conservative. For users with large numbers of friends, an attacker could take advantage of information that would appear on additional pages.

There are some limitations of our analysis. We only examined one university network at Facebook. Other networks may have different characteristics. For example, it may be harder to map names to email addresses on other networks. There could be sampling errors as well. In some ways, our results could be conservative. For example, attackers may be able to infer friend relationships by parsing the wall posts even if the friend attribute is hidden.

DEFENSE OPTIONS FOR USERS AND FOR SOCIAL NETWORKING WEBSITES

Given that the nature of social networks requires sharing information, it is difficult to counter the attacks that we outline in the previous section because they exploit the very nature of social networking. Because of this, there is a natural trade-off between the security and utility. We will discuss several possible defenses against this sort of attack here, along with their potential usability impact.

User-Centered Defense Strategies

For a Facebook user acting alone, defense options are limited. One could make his or her profile accessible only to friends. However, the user would still be susceptible to the birthday-invitation attack as well as the Facebook-template attack (a message that claims to be from Facebook about an update to a friend's page). The reason is that the user's friends may continue to have open profiles that expose the friends-relationship. For the user's defense to work, *user's friends would also have to make their profiles private*, which may be much harder to achieve, as it requires everyone in the group to consider risks and act on them collectively. This suggests the need for new mechanisms in social networking that help promote coordinated privacy policies. For example, perhaps there could be a feature in social networks that allows users to restrict their friends to those who have relatively safe privacy policies, or, only allow the friend relationship to be displayed on profiles of friends with safe privacy policies.

Another interesting defense strategy could be for Facebook (in collaboration with a school) to maintain a fake network of profiles as members of school networks whose sole purpose is to help provide an early warning of context-aware spam. The idea is similar to the deployment of honeypots in computer networks. If the attacker chooses to target a fake profile, Facebook and the school will get an early sample of the emails and can warn its users of the danger.

Social Network Interface Design and Privacy Policies

Most of our attacks required us to be able to map the first and last names from users' profiles to their email addresses by looking up an external directory. For Facebook, one possible solution would be to make the directory lookup less effective by removing the last name from profiles. This assumes that first names are much more common and thus less likely to lead to the right hit on a directory lookup. Unfortunately, removing last names would be detrimental to usability, as it would eliminate the ability of users to disambiguate others who share the same first name. This problem could be addressed by supporting a query to confirm last names, but this really offers no protection because candidate full names could be found by the attacker from the external directory and queried (note that most university directories allow lookups on first names and will return candidate choices). We judged this solution to only hurt usability and not provide significant protection.

Another defense strategy is to use images instead of text for sensitive attributes that can be part of vectors for an attack. Facebook already does this with email addresses. This could also be done for names of friends of the user. These associations are even useful for targeting people with closed profiles. However, just converting the names of friends to images is insufficient. Firstly, it only helps users with closed profiles, because the html link to an open profile is enough to get the user's full name anyway. A more effective strategy would be to make *all* names on profiles appear as

images, including the owner's. Of course, this strategy, while raising the bar, could be overcome by a determined attacker who has access to high quality OCR software or is able to use human computations to extract text from the images [22].

One extreme solution is to block most contextual information on profiles from being accessed by non-friends. This prevents any sort of association being made by an attacker outside of the circle of friends. (This appears to be the strategy adopted by LinkedIn.) A new user is still able to find his or her friends by searching for them explicitly, but it reduces the contextual information available to the attacker. Such an option would limit the scope of attacks. Unfortunately, this solution also makes it very difficult to find newly met acquaintances if, say, only a first name is known, so there are some functionality trade-offs. It potentially risks making the network "less interesting" to regular users, since it is more difficult to browse the network and find interesting information about friends of friends or acquaintances who are not yet friends.

A less extreme approach, but also less safe, is as follows. Suppose that person A has an open profile. As a matter of policy, full names and links to A's friends' profiles are only shown to A's friends. As a result, an attacker would be unable to find a "sender" email address to forge. (The email address, we assume, can only be reliably acquired from a Facebook profile or by searching with a full or part of a name at another source such as a directory.) The approach would be better than the current system, though some risks would still exist if profiles contain full names of the owner. Friendships often occur in clusters. Attackers could use heuristics to map initials to full names by clustering users who have a lot of common friends, identified with similar set of initials, and then examining their profiles.

Another defense we recommend is to make it harder for the attacker to acquire fake accounts on a social network. Facebook limits the number of "networks" that any Facebook profile can be a part of. However, it enforces no limit on how many Facebook profiles a person can have, although each Facebook profile requires a unique email address. Unfortunately, email addresses that are acceptable to Facebook can be easily forged by simply compromising one machine on a campus network and setting up email accounts and an email server on that machine. The job of the attacker could be made harder by restricting email addresses to top-level domain addresses of a university, which are often harder to forge (e.g., person@univname.edu rather than person@machine. univname.edu).

Facebook, at the time of this study, exposes the date of birth on open profiles by a default. The date of birth is required at the time of registration (presumably to validate that the person is older than 13 years). We recommend tightening the policy so that, by default, birthdays are not exposed on a profile, except to friends. Unfortunately, this may only help in reducing the effectiveness of the attack, and not completely eliminate it. Attackers could potentially scrape birthdays from the wall posts on a person's profile because they often contain birthday greetings from friends.

We are aware that Facebook has recently provided more fine-grained privacy controls. It is an open question as to whether the lack of finer-grained options was preventing users from better-controlling their privacy. There is a tradeoff – users may find advanced controls more complicated to use, whereas previously there were just a few. One solution may be to choose better privacy policy defaults based on the results from this study so that contextual information is generally available only to friends. User-interface design should also be improved to present contextual information (especially data that links users to each other) in such a way that it is harder to extract with automated tools.

Defense Techniques from Other Social Networking Sites

We now briefly examine the design of other social networks, such as MySpace and LinkedIn, which could help suggest defense strategies.

LinkedIn

In LinkedIn, most users' profiles are only visible to other users directly connected to that user in the network. In principle, this should make LinkedIn much more immune to the types of attacks we discussed in this paper. LinkedIn does provide operations to look up members and find the distance to them, including the name of a friend who is closer to the member. This could allow one to discover nodes that are at a distance two away. Nevertheless, it appears that LinkedIn does make it harder to discover the friend-friend relationships for users, though a more detailed analysis would be required to confirm this fact. Of course, as discussed earlier, this does reduce the ability of honest non-friends to browse the network and establish new relationships.

MySpace

Our analysis indicates that MySpace profiles tend to be much more customizable than Facebook's, with every user laying out their page in different ways. While MySpace profiles do tend to contain large amounts of personal data, it is unorganized and unlabeled. This could make it more difficult than on Facebook for attackers to write automated scripts that extract meaningful information. Another barrier, as noted in an earlier study, is that MySpace users are more likely to enter false information on their profiles. This could help reduce the scope of some attacks. For example, if the birthday on a profile is incorrect, then emails that use a birthday as the context may be less believable. Finally, it could be more challenging for attackers to link a profile with an email because users sometimes supply incorrect names on MySpace (33% of the time, see [6]). Of course, none of the above characteristics of MySpace are likely to be attractive defense options for Facebook as they all

potentially detract from usability or reduce the accuracy of content for all of Facebook's users.

CONCLUSION

In this paper, we analyzed Facebook, a popular social network among high school and university students, for its degree of vulnerability to different kinds of context-aware attack email. Our goal was to understand aspects of Facebook's policies and usage that may make its users vulnerable to sophisticated attacks via context-aware email. We classified attacks into three types: relationship-based attacks, unshared-attribute attacks, and shared-attribute attacks. We analyzed over 7000 randomly-accessed profiles on one university social network. By combining data from public profiles at Facebook with data from publicly accessible university directory services, we estimated that close to 85% of users could be accurately targeted with sophisticated context-aware attack email. Furthermore, even users with strict privacy policies on their profiles are almost equally vulnerable. We presented a detailed analysis of the potential risks for different types of attacks. We then discussed some potential defenses, along with trade-offs, which could help to reduce the risks. Our findings indicate that systematic defenses, which require changes at Facebook, are more likely to work than individual defenses.

REFERENCES

1. Arrington, M. *85% of College Students use Facebook*, Sept. 2005. (http://www.techcrunch.com/2005/09/07/85-of-college-students-use-facebook/)
2. Barabási, A., Albert, R. and Jeong, H. Scale-free characteristics of random networks: the topology of the world-wide web, *Physica A* 281 (2000), 69-77.
3. boyd, d. m. Friendster and publicly articulated social networking. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems,* ACM Press (2004), 1279-1282.
4. Brodkin, J. Phishing researcher 'targets' the unsuspecting, *Network World*, 24, 31 (Aug. 2007), 26.
5. CBC News, *Facebook 'ideal' for phishing attacks: researcher*, April 2007. http://www.cbc.ca/technology/story/2007/04/13/tech-facebookphishing-20070413.html
6. Dwyer, C., Hiltz, S., and Passerini, K.Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace, *Proc. 13th Americas Conf. Information Systems*, Association for Information Systems, 2007.
7. ESPC/Ipsos, Email Survey Summary, December 2006. http://www.espcoalition.org/ESPC_Ipsos_Survey_Executive_Summary.pdf
8. Furnell, S. and Ward, J. Malware comes of age: The arrival of the true computer parasite*, Network Security*, 2004, 10 (October 2004), 11-15.
9. Gross, R. and Acquisti, A. Information revelation and privacy in online social networks (the Facebook case). *ACM Workshop on Privacy in Electronic Society (WPES),* ACM Press (2005), 71-80.
10. Hodge, M. The Fourth Amendment and Privacy Issues on the "New" Internet: Facebook.com and Myspace.com, *Southern Illinois University Law Journal*, Fall 2006.
11. Jackson, M. O. A Survey of Models of Network Formation: Stability and Efficiency, in *Group Formation in Economics; Networks, Clubs and Coalitions*, edited by Gabrielle Demange and Myrna Wooders, Cambridge University Press: Cambridge U.K., 2004.
12. Jackson, M.O. and Rogers, B.W. Meeting strangers and friends of friends: How random are social networks? *American Economic Review 97* (2007), 890-915.
13. Jagatic, T., Johnson, N., Jakobsson, M. and Menczer,, F. Social Phishing, *Comm. ACM,* 50, 10, (Oct. 2007).
14. Jakobsson, M. and Ratkiewicz, J. Designing ethical phishing experiments: a study of (rot13) ronl query features. In Proc. WWW '06, ACM (2006), 513–522.
15. Jones, H. and Soltren, J.H. *Facebook: Threats to Privacy*, MIT manuscript, December 2005. Available at http://www.swiss.ai.mit.edu/6095/student-papers/fall05-papers/facebook.pdf.
16. Lampe, C., Ellison, N., and Steinfield, C.A face(book) in the crowd: social Searching vs. social browsing, In *Proc. CSCW*, ACM Press (2006), 167-170.
17. Lampe, C. A., Ellison, N., and Steinfield, C. A familiar face(book): profile elements as signals in an online social network. In *Proc. CHI'07*, ACM Press (2007). 435-444.
18. Liam Tung, *Social networking 'addiction' aids phishing*, May 2007. http://www.zdnetasia.com/news/security/0,39044215,62027706,00.htm
19. Newman, M.E.J., Forrest, S., and Balthrop, J. Email networks and spread of computer viruses, *Physical Review E* 66, 035101(R) (2002), 1-4.
20. Symantec, *Report: Hackers Turning to Social-Networking Sites*, September 2006.
21. Tsow, A., and Jakobsson, M. *Deceit and Deception: A Large User Study of Phishing,* Technical Report TR649, Indiana Universit*y*, August 2007.
22. Von Ahn, L. and Dabbish, L. Labeling images with a computer game. In *Proc.* CHI '04. ACM Press(2004), 319-326.