

Technical Report: November 16, 2012

The High-School Profiling Attack: How Online Privacy Laws Can Actually Increase Minors' Risk

Ratan Dey
Polytechnic Institute of
New York University
Email: ratan@cis.poly.edu

Yuan Ding
Polytechnic Institute of
New York University
Email: dingyuan1987@gmail.com

Keith W. Ross
Polytechnic Institute of
New York University
Email: ross@poly.edu

Abstract—Lawmakers, children’s advocacy groups and modern society at large recognize the importance of protecting the Internet privacy of minors (under 18 years of age). Online Social Networks, in particular, take precautions to prevent third parties from using their services to discover and profile minors. These precautions include banning young children from joining, not listing minors when searching for users by high school or city, and displaying only minimal information in registered minors’ public profiles, no matter how they configure their privacy settings.

In this paper we show how an attacker, with modest crawling and computational resources, and employing simple data mining heuristics, can circumvent these precautions and create extensive profiles of tens of thousands of minors in a targeted geographical area. In particular, using Facebook and for a given target high school, we construct an attack that finds most of the students in the school, and for each discovered student infers a profile that includes significantly more information than is available in a registered minor’s public profile. An attacker could use such profiles for many nefarious purposes, including selling the profiles to data brokers, large-scale automated spear-phishing attacks on minors, as well as physical safety attacks such as stalking, kidnapping and arranging meetings for sexual abuse.

Ironically, the Children’s Online Privacy Protection Act (COPPA), a law designed to protect the privacy of children, indirectly facilitates the attack. In order to bypass restrictions put in place due to the COPPA law, some children lie about their ages when registering, which not only increases the exposure for themselves but also for their non-lying friends. Our analysis strongly suggests there would be significantly less privacy leakage in a world without the COPPA law.

I. INTRODUCTION

It is generally recognized that protecting the Internet privacy of minors (under 18 years of age in the US) is important, with modern society manifesting this concern in many ways. The US government, through the Children’s Online Privacy Protection Act (COPPA) [1], requires commercial Web sites to obtain affirmative consent from parents before

children under 13 can create an account. Many consumer, privacy and child advocacy groups continue to actively lobby governments to provide better privacy protection for minors [2]. The US Congress is currently considering new bills to strengthen online safeguards for children and teens [3], [4],[5].

Online Social Networks (OSNs) additionally take measures to protect the privacy of minors. Facebook, for example, treats minors and adults with distinctly different policies related to information sharing, how members can find each other, and how they can contact each other [6]. Facebook currently bans young children (under 13) from joining, does not list minors when searching for users by high school or city, and displays only minimal information in registered minors’ public profiles, no matter how they configure their privacy settings.

In this paper we show how a third party, with modest crawling and computational resources, and employing simple data mining heuristics, can circumvent these precautions and create extensive profiles of tens of thousands of minors in a targeted geographical area. In particular, using Facebook and for a given target high school, we construct an attack which finds most of the students in the school, and for each discovered student infers a profile which includes significantly more information than is available in a registered minor’s public profile. The additional information minimally includes, for each discovered student, the student’s current high-school, graduation year, inferred birth year, and list of school friends. The generated profiles of about half of the identified minors also include varying amounts of additional information, including shared photos and wall postings. The information is collected *passively*, that is, without attempting to establish friend links with any of the students. As discussed in Section II, an attacker could use such profiles for many nefarious purposes, including selling the profiles to data brokers, large-

scale automated spear-phishing attacks on minors, as well as physical safety attacks such as prospecting candidate children for stalking, kidnapping and arranging meetings for sexual abuse.

Using off-line channels, it is difficult for an attacker to obtain complete lists of students attending a given target school. For example, in the course of the research for this paper, while seeking ground-truth data, we contacted administrators of four high schools and asked them to provide us with a list of names of all students currently attending their schools, with assurances of keeping the lists entirely confidential as well as not mentioning the names of the schools in this study. But the administrations of these high schools would not provide the lists, even with such assurances, fearing potential lawsuits from parents or other legal actions. High-school websites today also do not publicly provide lists of current students.

It is also difficult for an attacker to obtain complete lists of students attending a given target school directly from OSNs. As of September 2012, and documented in this paper, Facebook takes explicit measures to prevent people from obtaining school lists directly from its site. Although Facebook allows its members to search for other members who are associated with any given high school or city, *the search results returned by the service do not include registered minors*; for a high school search, they only include members who are registered as currently being 18 years or older, with the vast majority of the results being alumni of the high school. Because of this measure, it is not possible for an attacker to *directly* use Facebook’s search service to collect the names of the students at any target high school and attempt to profile them. Google+ also takes similar measures to protect the privacy of minors, as described in the body of this paper.

Ironically, the privacy leakages described in this paper are indirectly exacerbated by the COPPA law, which was designed to protect minors’ privacy. Given economic costs, social concerns, and technical issues, most online services — including Facebook and Google+ — choose to avoid the COPPA obligations by banning users younger than 13. Upon creating an account, these sites ask users for their birth date to determine if they are 13 or older. If the user indicates being under 13 years of age, the site prevents the user from creating an account. In response to this restriction, many underage users lie about their age to gain access to online social networks [7] [8]. (In fact, parents are often complicit in helping their children join the OSN [9].) For example, in order to gain access to Facebook, an 11-year-old boy may say he is 13 years old or may even say he is over 18. Facebook will then consider this boy older than he actually is, and eventually consider him an adult when his registered age becomes 18, even though he’s actually still a minor. Therefore, when searching for users by high school, although Facebook only returns members who are registered

adults, a small fraction of these registered adults are in truth minors. By identifying the minors returned by the search results, and performing statistical processing on their friend lists, we show it is possible to discover most of the students in the target high school and, for each discovered student, create a profile that contains significantly more information than should be available in a minor’s public profile. Thus, the COPPA law has inadvertently set the stage for widespread inference of minors’ private information.

To demonstrate the feasibility of the high-school profiling attack, we carried it out on one small private high school and two relatively large public high schools, located in different geographical regions in the USA. Our institution provided us with an IRB to perform the research under the condition that we keep private all collected and inferred information about individuals and only release aggregated results. For the smaller high school, we were able to obtain, through a confidential off-line channel, ground-truth information including the names of all the students in the high school and their graduating classes. For the larger high schools, we were able to obtain limited ground-truth information for a small set of students by scraping Facebook. This ground-truth information enables us to validate the approach and measure the success of the attack.

We estimate how much privacy leakage would occur in a world without the COPPA law and compare the estimate to the extent of leakage in our current world with COPPA. Our results suggest that an attacker not only can discover more minors, but can also build more extensive profiles than what would be the case in a world without COPPA. Thus, in terms of third-party privacy attacks, COPPA actually puts minors at greater risk than they would be if the law had never been enacted.

To our knowledge, this is the first paper that (i) identifies the third-party privacy problem in OSNs for minors, (ii) quantifies the extent of the privacy leakage, and (iii) investigates the impact of a law on privacy leakage. This paper should hopefully increase awareness about this class of attacks, some of which are possibly already being carried out by data brokers today (although we currently have no evidence to support this). As part of responsible disclosure, we informed both Facebook and Google about the attack.

This paper is organized as follows. In Section II we briefly outline the consequential threats resulting from the high-school profiling attack. In Section III we provide definitions of terms used throughout the paper, discuss Facebook’s and Google+’s policy for minors, and discuss ethical issues associated with the collection and analysis of our data. We present the details of the high-school profiling attack in Section IV. We evaluate the success of the attack for three high schools in Section V. In Section VI we investigate to what degree minors can be profiled using the attack and in Section VII we show several possible ways to contact minors. In Section VIII we estimate the extent of leakage that

would occur in a world without COPPA. In Section IX we consider one promising countermeasure, namely, disabling reverse lookup. In Section X we discuss relevant prior work.

II. CONSEQUENTIAL THREATS

Suppose that an attacker, with modest crawling and computational resources, for a given target high school, is able to determine all the students in the school and profile them, with the profiles containing a varying amount of information, but minimally including full name, profile picture, gender, high school name, graduation year (i.e., grade), and high-school friends. (The attacker could further infer current city from the high school and birth year from the graduation year.) We call these profiles the *high-school profiles*. Moreover, suppose the attacker has a means to send messages directly to many of the students, and can send friend requests to all of the students. We now describe some of the consequential threats.

The first major threat is that of data brokers collecting high-school profiles and selling them to others, such as advertisers, college recruiters, and employment agencies. Because the teen market surpasses US\$200 billion in the US alone, it is not surprising that data brokers are already seeking to compile dossiers on children [10] [11]. When collecting the high-school profiles, the data broker may first choose to target high schools in wealthy communities, since those profiles may be more valuable to marketers.

By leveraging the information in the high-school profiles, data brokers can also enhance the profiles by linking them with other personal data available online and from public records. For example, by obtaining voter registration records (which most states make available for a small fee), the data broker can use the last name and city in the high-school profiles to link the students to parents in the voter registration records, thereby determining the street address of many of the students. For those students with friend lists in the high-school profile, if a parent appears in the friend list, then the street-address association can be done with greater certainty. As another example, for many students, the first name, last name and city in the high-school profiles can be linked with Skype profile information in the Skype directory, thereby augmenting the profile with a means of calling and videoconferencing with the teenager.

The second major threat is that of a pedophile, who seeks to use the Internet to arrange sexual encounters with children. A cursory Internet search reveals that such occurrences are widespread today; for example, recently a man allegedly used Facebook to arrange meetings and have indecent contact with seven different girls, ranging in age from 13 to 15. The district attorney for the case stressed the importance of minors “not sharing personal information online, like full names, ages, addresses, phone numbers and school information” [12]). A pedophile could launch the high-school profiling attack himself, using the acquired profiles to prospect for victims. As a first step, the attacker could use the profiles

to narrow down the candidates in the target community. The attacker could then leverage the profile information to perform social engineering attacks and establish online contact with the candidates. Similarly, a kidnapper, who might target schools in wealthy communities, can use the profiles to discover, narrow down, and contact prospects.

Finally, the profiles could also be used to fuel a large-scale and highly personalized spear-phishing attacks against minors. Messages could automatically be generated which mention the target students’ high schools, graduation years, and friends, tricking the targets into installing malware on the family computer, for example.

III. PRELIMINARIES

Throughout this paper we define a *minor* to be any person who is currently under 18 years old. Anyone 18 years or older is said to be an *adult*. Note that most students currently attending a high school are minors. (A fraction of the final-year students may be adults, with the fraction increasing each month in the school year.) OSNs typically require users to specify their birth date (day, month, and year) when they register. As discussed in the Introduction, some users may lie about their birth dates when creating accounts in order to circumvent the minimum age requirement. A user is said to be a *registered minor* if the OSN believes the user is currently a minor based on the registered birth date. We define a *registered adult* in a similar manner. In the context of Facebook, we say a user (say, Alice) is a *stranger* to another user (say, Bob) if all the following conditions are satisfied: (i) Alice is not a friend of Bob; (ii) Alice is not a friend of friend of Bob (that is, Alice and Bob have no mutual friends); and (iii) Alice does not belong to any of Bob’s school or work networks.

A. Facebook and Registered Minors

In Facebook, registered minors have a different experience with privacy than do registered adults. We now highlight the differences that are relevant to the current study. Table I shows the information about a user available to a stranger for when the user keeps the default settings and for when the user configures the setting for maximum sharing (worst case). A check in the box means the information is available to the stranger for the specific scenario. As shown in Table I, when a stranger visits a registered minor’s profile page, only a limited amount of information is available to the stranger: at most the user’s name, profile photo, networks joined, and gender are available. (Typically less, depending on how the user configured her privacy settings. For example, typically less than 10% of registered minors specify network.) Further, the “Message” button will never be visible to a stranger.

We say that *only minimal information is available about a user* (registered minor or adult) if a stranger, when visiting the user’s public profile, sees at most name, profile photo, networks joined, and gender, and the “Message” button is not available. It follows that if a stranger visits a user’s public

profile and more than the minimal information is available, then the user must be a registered adult. If only minimal information is available, then the user is very possibly a registered minor, but may instead be a registered adult who configured her privacy settings to show minimal information in her public profile.

OSNs typically provide a friend-search feature, allowing its users to find new friends from different parts of their past and current lives, including friends from previous high schools. Facebook provides this feature in its “Find Friends Portal” [13], where a user can search for potential friends by inputting either hometown, current city, high school, mutual friend, college or university, employer, or graduate school. When a stranger does a high school search by the high school name, Facebook returns a few hundred users who are associated with the target high school. The stranger can also attempt to obtain additional users by creating additional fake accounts. We wrote a script that collects users in this manner. The script takes as input the target high school’s Facebook ID, a username and password for a fake account, and outputs several hundred unique Facebook user IDs.

We observed in the course of experiments that *Facebook does not return any registered minors when a stranger searches with the Find Friends Portal*. We verified this claim by carrying out an experiment with a high school for which we have the complete list of current students at the high school, as well as the complete list of recent alumni. Using multiple stranger accounts, we obtained from the Find Friends Portal a total of 352 unique Facebook users associated with the target high school. We then checked to see if any of the users are currently students at the high school by matching the 352 Facebook users with the ground-truth lists. Although, a small fraction of the 352 users were found to be current students in the target high school, all of those students make available more than minimal profile information, and are therefore registered adults (some which who recently turned 18 and others who are minors who lied about their birth dates during registration). Thus, none of 352 Facebook users returned by the Find Friends Portal were registered minors at the time of the experiment (June 2012).

In summary, in an attempt to act responsibly towards minors, Facebook takes some precautions to protect minors’ privacy. We observed and verified that Facebook does not return registered minors when a stranger searches by high school. Also, when a stranger visits a registered minor’s public profile page, only limited information is made available, no matter how the minor configures the privacy settings. In particular, a minor’s high school, graduation year, and friend list are never directly available to a stranger.

B. Google+ and Registered Minors

Although the focus of this paper is on Facebook, we briefly mention here that Google+ is also susceptible to the high-school student profiling attack. Like Facebook, to create a

Google+ account, the user must register as 13 years or older [14] [15]. Google+ also provides a mechanism for searching for users associated with a high school [16].

Unlike Facebook, Google+, which uses circles, has asymmetric friendship links. For example, for Alice there is one set of users in her circle; and there is a second set of users who include her in their circles. Google+ also provides various safety guidelines for teens [17]. Google+ registered minors also have different default privacy settings than do registered adults, as shown in Table II.

C. Legal and Ethical Considerations

To perform the research described in this paper, we implemented customized crawlers that visit public Web pages in Facebook and download the HTML source code of each Web page. Our parser then extracted relevant data from the HTML source code and stored the data in an SQL database.

Crawling data in OSNs is an ethically sensitive issue. One question that arises is if it is ethically acceptable and justifiable to conduct crawling experiments in social networks? We believe that the only way to reliably estimate success rates of attacks in the real-world is to use realistic experiments. We nevertheless took several precautions while crawling. First, we only accessed user information that was publicly available. Second, by implementing sleeping functions and limiting our study to three high schools, the crawling was not particularly aggressive and didn’t perturb the performance of Facebook.

We also obtained IRB approval for this work from our university. As part of responsible disclosure, we informed both Facebook and Google about the attack. In October 2012. Because of the sensitive nature of the information we gathered and inferred, we will not be making our data sets public and we will not explicitly identify the high schools involved.

IV. THE HIGH SCHOOL PROFILING ATTACK

We now describe our basic version of the high-school profiling attack. The attacker begins by selecting a target high school. Let M be the set of all the students currently attending the target high school with active accounts in the OSN. The goal of the attack is to find most of the students in M and obtain (or infer) as much profile information as possible about each of those students.

A. Threat Model

We do not require the attacker to be an OSN friend, or a friend-of-a-friend, of any of the students in M , that is, the attacker may be a stranger to all the students in the high school *throughout the duration of the attack*. With sufficient computational resources, the attack could therefore be launched against hundreds or even thousands of high schools.

We assume that there is a means for an attacker to find the OSN IDs of a small set of users who are *both* current students

TABLE I
FACEBOOK: DEFAULT AND WORST-CASE INFORMATION AVAILABLE TO STRANGERS

	Default Availability for registered minors	Default Availability for Registered Adults	Worst-case for Registered Minors	Worst-case for Registered Adults
Name, Gender, Networks, Profile Photo	✓	✓	✓	✓
HS, Relationship, Interested In		✓		✓
Birthday				✓
Hometown, Current City, Friendlist		✓		✓
Photos		✓		✓
Contact Information				✓
Public Search		✓		✓

TABLE II
GOOGLE+: DEFAULT AND WORST-CASE INFORMATION AVAILABLE TO STRANGERS

	Default Availability for Registered Minors	Default Availability for Registered Adults	Worst-case for Registered Minors	Worst-case for Registered Adults
Name, Profile Picture	✓	✓	✓	✓
Gender, Employment, HS, Hometown, Current City		✓	✓	✓
Home and Work Phone			✓	✓
Relationship, Looking			✓	✓
Birthday			✓	✓
Photos			✓	✓
Public Search		✓	✓	✓
In Your Circles		✓	✓	✓
Have You in Circles		✓	✓	✓

in the target high school *and* registered adults. Moreover, we assume that this set includes several students in each of the graduation years (typically four years in the US).

One way this assumption can be satisfied is that the OSN provides a search function which allows a user to input a target high-school name, and the search function returns a more-or-less random subset of registered adults attending, or having previously attended, the school. OSNs — including Facebook and Google+ — typically provide such a friend-search feature, allowing its users to find friends from different parts of their past and current lives, including friends from previous high schools. Importantly, the assumption does not require that the search function return users who are registered minors: *the search function may only return users registered as adults*, thereby taking some measures to protect minors from being easily discovered and profiled.

Another way this assumption can be satisfied is if the OSN has a friend recommendation system. The attacker can create a fake user with a profile stating he is a current student in the school. The friend recommendation system may then recommend registered adults who also indicate they are current students in the school.

B. The Basic Attack: Exploiting Lying Minors

For concreteness, we describe the attack in the context of the OSN providing a search function that allows a user to

input a target high-school name. For any user u in the OSN, let $F(u)$ be the user's current set of friends. For some users, $F(u)$ will be visible on the user's public profile; for other users $F(u)$ will not be publicly available. The attack in its most basic form operates as follows.

- 1) The attacker inputs the name of the target high school into the OSN's high-school search function. The search function returns a list of members who are associated with the target high school. The attacker may use a script to automatically scroll down the page (thereby sending additional HTTP requests with AJAX) in order to get a longer list of members. The attacker may also use multiple accounts when searching. We refer to the set of all the members found in this manner as the *seeds* and denote the set by S .
- 2) The attacker uses a crawler to download the public profile pages (but not friend lists) for each of the seeds, parses the pages, and determines the users who indicate they currently attend the target high school (by listing their high school as the target high school and providing a graduation year that is the current year or a future year). Let C' be the subset of seeds who explicitly indicate (in their public profiles) that they are currently students in the target high school. (Most of the users in C' will be minors who lied about their age during registration.) Let C be the subset of users in C'

who make their friend lists public. We refer to C as the *core set*. As we will see, the number of core users is typically fairly small, on the order of 5% of the number of students in the high school. For each user in set C , we know the user’s graduation class year. Assuming that the high school is a four-year school, denote C_1 , C_2 , C_3 , and C_4 , for students in the first second, third, and fourth school years in the core set C .

- 3) For each student $u \in C$, the attacker downloads the friend list, $F(u)$, from the OSN. Let K be the set of all friends obtained from the core users, that is,

$$K = \cup_{u \in C} F(u).$$

We refer to K as the *candidate set*. Our experiments show that the number of candidates will approximately be one order of magnitude greater than the target high school size.

- 4) We expect some of the users in K to be current students in the target high school. We now try to determine which ones. For each candidate $u \in K$, we use *reverse lookup* to determine its friends in the core. Specifically, for each $u \in K$, we determine the set of friends in the core set for each of the four graduation years:

$$G_i(u) = \{v \in C_i : u \in F_v\}, \quad i = 1, 2, 3, 4. \quad (1)$$

Clearly each $G_i(u) \subseteq F(u)$. Note that to obtain the $G_i(u)$ ’s, the attacker *does not* have to obtain the profile pages or friend lists of any of the users in the large candidate set K . In fact, user u ’s friend list may not even be directly available to strangers.

- 5) For each candidate $u \in K$, the attacker calculates the fraction of users in each of the core class sets with whom the candidate is friends, and then calculates the maximum of these four fractions. Specifically, the attacker calculates

$$x(u) = \max_{1 \leq i \leq 4} \frac{|G_i(u)|}{|C_i|} \quad (2)$$

- 6) The attacker rank orders the users in K according to their $x(u)$ values, from highest to lowest. The attacker chooses a threshold t in the vicinity of the total number of students attending the high school (which can typically be found from Wikipedia or some other source). The attacker then considers the first t students as current students in the target high school (as well as the students in the set C'). Let T denote the set of t students and $H = T \cup C'$. The attacker also classifies each such student $u \in T$ into a graduating year according to the highest $|G_i(u)|/|C_i|$ value, $i = 1, 2, 3, 4$.

At the end of these steps, the attacker has a set of OSN users H believed to be students at the target high school. The attacker has also classified all the students in H by graduation class year. If the attacker seeks additional profile information,

he can proceed to download the public profile information and public friend lists (when available) for the students in T . The attacker then uses this information along with statistical inference to extend the profiles of the students in H , as we will discuss in Section 6.

Note that the attack relies on the attacker’s ability to obtain a small set of core users, that is, finding a set of users for whom the attacker knows with certainty that the users are in the high school and knows their graduation year. Because the search function may not return any registered minors, *a priori* the core set will have no students in years 1 to 3 and few in year 4. However, because a significant fraction of minors lie about their birth dates when creating accounts [7] [8] [9], in order to circumvent the age restriction due to the COPPA law, it is indeed possible to obtain a core set from the search function including students distributed across the four years. Also note that the active is passive, that is, without attempting to establish friend links with any of the students.

C. Attack Performance

The set H , and the classification of its members by graduation year, is obtained by statistical inference and therefore may contain errors. For example, some of the users in H may be false positives, that is, they are not current students at the target high school. Furthermore, H may not contain all of the students in M . Two important measures for the performance of the attack are the *fraction of students from M found*, given by $|H \cap M|/|M|$, and the *number of false positives*, given by $|H - M|$. Note that by varying the value of the threshold t the attacker can trade off these two performance measures: increasing t should increase the fraction of students found but should also increase the number of false positives. In this paper we estimate these measures for each of are three test high schools.

D. Enhanced attack

We now describe an important enhancement of the attack, which requires a relatively small amount of additional crawling. In the *enhanced attack*, after rank ordering the $x(u)$ ’s and selecting a threshold t , we download the public profile pages of the first $t(1 + \epsilon)$ users. (In this paper, we use $\epsilon = 1$ throughout.) Denote this set of users by $T+$. For each user u in $T+$, we then check the user’s profile to see if he indicates he is currently a student in the target high school. If so, we move u from $T+$ to C , thereby increasing the size of the core set. After doing this for all $u \in T+$, we recalculate $G_i(u)$ for each $u \in T+$ and $i = 1, 2, 3, 4$, and proceed from Step 5 in the Basic Approach.

The basic and enhanced approaches just described are natural and intuitive heuristics. In addition to these approaches, there are many possible heuristics one may construe based on the $G_i(u)$ data. It is also possible to explore traditional machine learning approaches. At the expense of doing additional crawling, it is also possible to collect more data

and develop refined inference approaches. For example, by crawling all the friend lists in the large candidate set K , we can determine many of the friendship relationships among the candidates, which can potentially be used to improve the statistical inference. As the purpose of our research is to demonstrate the feasibility of the attack rather than fully optimize it, we do not pursue these optimizations here.

E. Filtering

In order to possibly improve the performance of the basic and enhanced attacks, we also examine filtering out some of the candidate users. This filtering variation, as with the enhanced attack, requires that the attacker download the public profiles of the first $(1 + \epsilon)t$ users in the candidate set. After downloading these profiles, the attacker applies filtering rules to eliminate candidates who are likely former students at the target high school (and have transferred out or have already graduated). We used the following filter rules:

- *Graduate School*: The candidate specifies a graduate school in the public profile page.
- *Different High School*: The candidate provides *one* high school and that high school is different from the target high school.
- *High school graduation year*: The candidate provides a high-school graduation year that is not in the current year or in the subsequent three years.
- *Current city*: The candidate provides a current city other than the city in which the high school resides.

F. Estimating the Crawling Effort

Most OSNs employ anti-crawling techniques to protect the data of their members and the performance of their sites. Typically, if a member behaves suspiciously (for example, if he tries to access an overly large amount of user profiles in a short amount of time), the member’s account will be temporarily, or permanently, disabled. Therefore another important measure is the crawling effort required to perform the attack.

For the Basic Attack, the crawling effort has three components: (i) the number of HTTP GETs sent to obtain the IDs of the seed users S (Note that with AJAX, multiple HTTP GETs may need to be sent to get the entire page.); (ii) the number of HTTP GETs sent to obtain the public profile pages of the seed users in S ; (iii) the number of HTTP GETs sent to obtain the friend lists of each of the core users (again sending multiple GETs via AJAX). The approximate number of HTTP GETs sent is therefore given by $A \cdot R + |S| + |C| \cdot f/p$, where A is the number of accounts used, R is the number of HTTP GETs sent per account when gathering the seed list, f is the average number of friends a student has, and p is the number of friends gathered with a single HTTP request. (Currently, Facebook uses $p = 20$).

For the enhanced attack, we additionally (i) download the profile pages of an additional $(1 + \epsilon)t$ users, where t

is roughly the number of students in the target school, and (ii) download the friend lists for the augmented core set. In Section 5 we will show that the total number of requests for a typical school is small for both the basic and enhanced attacks.

V. RESULTS FOR THREE HIGH SCHOOLS

A. Data Sets

In order to estimate the success of the attack, we applied it to three US high schools, which we refer to as HS1, HS2, and HS3. We collected the data for HS1, HS2, and HS3 in March 2012, June 2012, and June 2012, respectively. HS1 is a small private urban high school with about 360 students. For this high school, we were able to obtain, through a confidential channel outside of Facebook, the complete student lists (segmented by graduation year) for the high school, and also complete alumni lists for recent graduation years. These lists enable us to evaluate the success of the attack. HS1 has a relatively high churn rate, with 10-20% of the students transferring in and out of the high school every year. Because of the high churn rate, it is a challenging problem to determine an accurate estimate of the current snapshot of the student body. However, we will see that even with this high churn rate, the basic attack provides good results.

For the HS1 students in the 2012, 2013, 2014, and 2015 graduating classes, we were able to find the Facebook IDs and public profile pages for $|M| = 325$ students. We did this essentially by running the basic attack on HS1, finding the users who were ranked the highest, and checking for their names in the ground truth list. We were not able to find the Facebook IDs for about 10% of the student body at HS1. Most of these remaining students most likely do not have Facebook accounts. A small number of them may have accounts with alias names that we could not match to the ground-truth list. The 325 students are roughly evenly distributed over the four years; for 112 students (34%) their friend lists are publicly available.

HS2 is a public suburban high school on the East Coast with a much larger student body of approximately 1,500 students. The school has diverse economic and racial demographics, with about 15% of the students being African-American, 10% Asian, and 10% Hispanic. HS3 is a public high school in a small city in the Midwest, also with approximately 1,500 students. Although neither for HS2 nor HS3 were we able to obtain complete ground-truth information, we were able to evaluate the attack based on partial ground-truth information mined from Facebook.

B. Initial Seed Set

We obtained initial seed sets from Facebook’s Find Friend portal, using two accounts for the smaller HS1 and four accounts for each of the larger high schools HS2 and HS3. Table III provides a summary of the data collected for the

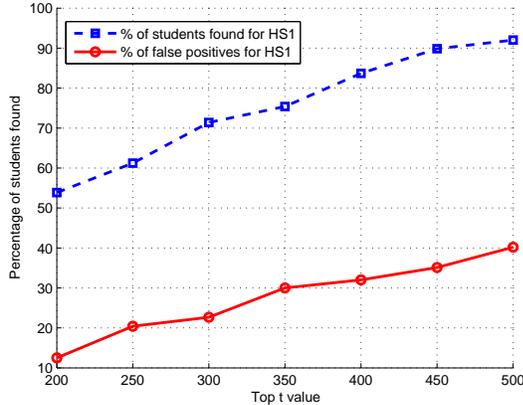


Fig. 1. Overall performance of enhanced attack for HS1

three schools. As shown in Table III, for HS1, HS2, and HS3, we found 18, 70, and 46 core users (with friend lists) and 6,282, 14,317, and 11,736 candidates, respectively. For the enhanced attack, we obtained 22, 152, and 178 (extended) core users for each of three high schools. For each high school, the number of core users is roughly 5% of the number of students in the school.

C. Crawling Effort

Table IV summarizes the approximate crawling efforts required to collect the data sets for the three high schools in Table IV. Note that the effort is quite small, with the number of HTTP requests sent being about twice the number of students in the target high school for the basic attack, and about five times the number of students in the target high school for the enhanced attack.

D. Results for HS1

Recall that for HS1 there are 325 students having Facebook accounts. Also recall that we have the complete ground-truth information for HS1 (i.e., the Facebook IDs and graduation years for all of the 325 students). The results for both the basic and enhanced attacks, with and without filtering, are shown in Table V for thresholds t ranging from 200 to 500. The set of users in each column includes the core users (or extended core users for the enhanced attack). In the notation x/y , x is the number of users from the set of 325 students that are found; and y is the number of users, from the set of x users, that are classified in the correct classification year. We see for the top 200, 300, and 400 cases, the enhanced attack with filtering gives the best results; for the top 500 case, the enhanced attack without filtering gives somewhat better results than the enhanced attack with filtering.

We see that the filtering indeed reduced the number of false positives for the threshold of top 200, top 300, and top 400 users. But for the larger threshold, the filtering actually increased the number of false positives. This can be explained as follows. On one hand, when we increase the threshold beyond 400, we add mostly false positives, since there are not many true positives remaining. On the otherhand, the filtering

also accidentally filters out some of the true positives, giving an overall decrease in performance.

As an example, let us suppose that the attacker decides to use the enhanced attack with filtering, and considers the top 400 users as students in HS1. Examining the column for 400 students in Table V, we see that with this choice of threshold, 272 (84%) of the 325 students are included in the attacker's set. So with this threshold, the attacker finds 84% of the high school student body (having Facebook accounts) with 128 false positives (32%). Moreover, of these 272 students, 250 (92%) have been classified in the correct graduation year. If the attacker wants to reduce the false positives, the attacker can declare only the top 200 users as students, in which case there are only 25 (13%) false positives, with 54% of the students found, of which 90% are classified in the correct graduation year. If the attacker can accept a larger number of false positives, he may instead choose the top 500 students, which would include 92% of the high school student body having Facebook accounts. We show these estimates for different choices of threshold t for the enhanced attack with filtering in Figure 1.

The results of obtaining 84% of the students in the high school, of which 92% are classified in the correct year, with 32% false positives are remarkable, particularly when considering the 10-15% annual churn rate at the high school. Many students attend HS1 for a short period of time. They make friends with the other students during their period of study, then their families move to another city. We manually inspected the 128 false positives (from the set of top 400 users) and found that about half of them were former students at HS1. For the other half of the false positives, they make very little public information available, so it is difficult to determine if they are former students or not (although most likely are since they have a large number of friends in HS1).

E. Results for HS2 and HS3

For each of the two large public high schools, in order to evaluate the performance of the basic and enhanced attacks, we collected a first set of seeds with four Facebook accounts and a second set of seeds with an additional four accounts. We use the first set of seeds to perform the attack; we use the second set for evaluation. Specifically, for HS2, from the second set of seeds we obtained 43 users who specify they are currently at HS2 and are not included in the first set of seeds. To evaluate the attack, we check to see which of these 43 test users are in our inferred set, and which of those are classified in the correct graduation year. For HS3, we obtained 47 such test users.

The results for HS2 for thresholds t ranging from 500 to 2000 are shown in Table VI. For example, for the enhanced attack with filtering, 36 out of the 43 test users were found within the top-1500 users; of those 36 test users, 35 were classified in the correct graduation year. We can see from Table VI that the enhanced attack with filtering gives equal

TABLE III
SEEDS, CORE USERS, AND CANDIDATES FOR THE THREE HIGH SCHOOLS

High school	# of students	# of students on Facebook	# of seeds	# of core users	# of candidates	# of extended core users
HS1	362	325	352	18	6,282	22
HS2	1,500 (approx)	N/A	1,559	70	14,317	152
HS3	1,500 (approx)	N/A	1,532	46	11,736	178

TABLE IV
CRAWLING EFFORT

	# of Facebook accounts used	# of HTTP requests to download seeds	# of profile pages downloaded	# of requests to download friend lists	total # requests for basic attack	total # of requests for enhanced attack
HS1	2	34	352	360	746	1,576
HS2	4	101	1,559	1,400	3,060	7,700
HS3	4	90	1,532	920	2,542	8,182

TABLE V
RESULTS FOR HS1 (WHICH HAS 325 FACEBOOK USERS)

	Top 200	Top 300	Top 400	Top 500
Basic attack without filtering	140/112	206/162	271/224	301/254
Basic attack with filtering	148/122	196/165	259/227	299/264
Enhanced attack without filtering	169/155	231/211	261/239	304/281
Enhanced attack with filtering	175/158	232/211	272/250	299/276

or better results than the other approaches. Note that unlike for Table V, the users in each top- t column in Table VI do not include the core (or extended-core) users.

Using the results in Table VI, we now estimate the total number of students found in HS2 for a given threshold t . To this end, let x_t be the number of test users found in the top t . For the basic attack, the set of actual high-school students discovered for a threshold t has two disjoint groups: (i) the core users; and (ii) the non-core high-school students who are discovered. The fraction of non-core high-school students who are discovered is given by $p = \text{non-core HS students discovered} / \text{non-core HS students}$. This fraction can be estimated by $(x_t / \# \text{ test users})$. Thus, an estimate of the number of students in the high school found with a threshold t is:

$$\# \text{ of core users} + \frac{x_t}{\# \text{ test users}} \times (\text{HS size} - \# \text{ of core users})$$

(For the enhanced attack we replace the number of core users with the number of extended core users.) To estimate the percentage of high-school students found for threshold t , we divide the above by the high-school size. To estimate the number of false positives for a threshold t , we use

$$t - \frac{x_t}{\# \text{ test users}} \times (\text{HS size} - \# \text{ of core users}),$$

since the false positives are those users among the top- t minus the expected number of students found in the the top- t (excluding the core users). To estimate the percentage of false positives for a threshold t , we divide the above by $\# \text{ of core users} + t$. We show these estimates for the enhanced attack with filtering in Figure 2. For example, for the top 1,652 users ($t = 1500$ plus the extended core users), the attacker can obtain 85% of all the HS2 students with 22%

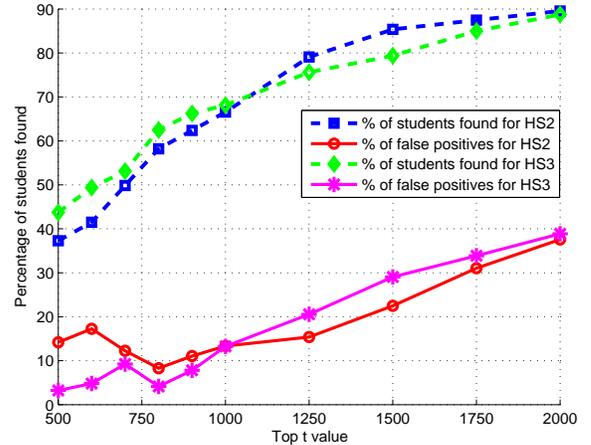


Fig. 2. Overall performance of Enhanced attack for HS2 and HS3

false positives in the set of 1,652 users. On the whole, the results in Figure 2 for HS2 are similar to those in 1 for HS1. Also,

The results for HS3 for thresholds t ranging from 500 to 2000 are shown in Table VII. Once again, the enhanced attack with filtering gives the best results. Figure 2 shows the performance of the attack for HS3.

F. Identifying Registered Adults who are Actually Minors

We now consider the problem of identifying registered adults who are almost certainly minors. An OSN could use this identification to tighten the privacy of the identified minors (e.g., display only minimal profiles and remove the “Message” button).

Almost all of the students in the first three years of high school should be minors. Some of these minors will be

TABLE VI
RESULTS FOR HS2 - 43 TEST USERS

Method Name	Top 500	Top 1000	Top 1500	Top 2000
Basic attack without filtering	13/12	25/24	32/29	36/31
Basic attack with filtering	13/12	25/24	32/29	36/31
Enhanced attack without filtering	13/12	26/25	35/34	37/36
Enhanced attack with filtering	13/12	27/26	36/35	38/37

TABLE VII
RESULTS FOR HS3 - 47 TEST USERS

Method Name	Top 500	Top 1000	Top 1500	Top 2000
Basic attack without filtering	18/17	26/21	29/22	34/25
Basic attack with filtering	18/17	26/21	31/22	36/26
Enhanced attack without filtering	15/15	29/27	34/30	39/35
Enhanced attack with filtering	17/17	30/27	36/32	41/36

minors registered as adults. We now identify the minors registered as adults for the ground-truth students in HS1, the top 1,652 users in HS2 and the top 1,678 users in HS3, limiting the identification to those users who have been classified to be in the first three years of high school. Due to Facebook’s privacy policy as discussed in Section III, a minor is unmistakably a registered adult, and is identified as such, if either public search is enabled or strangers can see any “public-profile adult attributes” such as relationship status, friend list, current city, hometown, interested in, educational information, employer information, or the “Message” button. Using this classification, we found 112 out of 238 students (47%), 700 of 1,226 students (57%), and 795 of 1,373 students (57%) at high schools HS1, HS2, and HS3, respectively, to be minors registered as adults. These numbers are roughly consistent with numbers obtained from surveys in [7] [8] [9].

G. Summary of Results

As discussed in Section 2, when using Facebook’s Find Friends Portal to search for users in a target high school, Facebook takes precautions to protect minors by not returning any registered minors. We have shown that an attacker, with relatively little crawling effort, can discover the majority of the students at the target high school. For example, we obtained 83%, 85% and 79% of all the students in HS1, HS2, and HS3, respectively, with false-positive rates of 32%, 22% and 29%. Moreover, for each high school student in the list, the attacker can determine the student’s graduation year with a high-level of accuracy. It may be possible to improve these results using more refined heuristics, machine learning, or iterative classification [18] [19] [20].

Furthermore, because we obtained consistent results from the three high schools, it appears that at least half of all US high school students lied about their age during Facebook registration. Although we did not perform the attack with Google+, based on the discussion in Section 2, the attack likely applies as well to Google+.

VI. EXTENDING THE PROFILES

Recall that when a stranger visits the Facebook page of a minor, in the philosophy of Facebook’s current privacy policy, the stranger should see minimal information, which at most includes the minor’s full name, profile photo, gender, and networks. However, due to statistical inference and many minors registering as adults, an attacker can leverage OSNs to significantly extend the profiles. In this section we quantify the amount of additional profile information is readily available to the attacker. We do this separately for two classes of minors: those who are registered minors, and those who are registered adults. Again, we do this for the users classified in the first three years of high school (since some of the fourth year students are adults).

A. Extending Profiles of Registered Minors

In addition to the minimal information (full name, profile photo, gender, and occasionally networks), we have shown that with high probability an attacker can infer current high school and graduation year. From the inferred high school, the attacker can also infer hometown and current city, and from graduation year the attacker can further estimate the minor’s birth year.

In performing a social engineering attack on a minor, in addition to having the minor’s minimal information plus high school and graduation year, the attacker would most likely want to know who are the minor’s friends. This information is not directly available to a stranger for a registered minor. But the attacker can use “reverse lookup” to obtain partial friend lists of the registered minors. Specifically, after obtaining a set H of (likely) current students at the target high school, the attacker downloads the friend lists for all users in H whose friend lists are publicly available. A student in H , say Alice, without a public friend list, will typically be in the friend lists of other students in H who make their friend lists public. With this information, the attacker can determine at least a portion of Alice’s friends. We applied reverse lookup to the registered minors in HS1 and the inferred registered minors in HS2 and HS3. For each of these minors we were able to create friend lists. In particular, we found on average 38, 141,

129 friends per registered minor in HS1, HS2, and HS3. (On average HS1 students have fewer high-school friends since it is a much smaller high school.)

In summary, for each registered minor discovered through the high-school profiling attack, the attacker can create a profile with full name, profile photo, gender, a large subset of friends, high school, graduation year, hometown, and an estimate of birth year. This is substantially more information than what Facebook makes publicly available, no matter how the registered minor configures her privacy settings. As discussed in Section II, this information can serve as a base for creating more comprehensive profiles, by matching the information with public records and other online sources.

B. Extending Profiles of Minors Registered as Adults

We now consider the additional information that an attacker can obtain for a minor who is registered as an adult. In this case, significantly more information is often directly available, depending on how the user has configured her privacy settings. Possible additional information that can be collected by a stranger — beyond the information an attacker can obtain and infer for a registered minor — includes shared photos, photo tags, full friend list, relationship info, interested in, wall postings, likes, favorites, political views, religious views, videos, links, website, birthday and contact info (such as personal email address, IM screen name, address, phone number). For HS1, HS2, and HS3 we determined how much additional information is available for some of these attributes. The results are shown in Table VIII.

A stranger can obtain a significant amount of information about a registered minor, as we saw in the previous subsection. But the stranger can often obtain even more information for a minor registered as an adult. For example, as shown in Table VIII, an attacker can obtain on average over 50 shared photos in the two large high schools, and has access to the “Message” link for more than 86% of the minors registered as adults in all three high schools.

VII. CONTACTING MINORS

Up to this point, the attacker has identified Facebook users who are likely to be current students in the target high school. There are several possible ways the attacker can send messages to minors. Such messages can contain attachments with malware, links to phishing sites, or ploys to meet in person.

1) *Contacting through Facebook’s message service:* We have seen that about half of the minors in high schools are registered adults. If such a minor doesn’t change her default settings, the “Message” link will appear on her public profile page. From Table VIII, we can see that for more than 86% of the minors registered as adults have the “Message” link enabled in all three high schools. The overall percentage of minors with the “Message” link enabled therefore in the 40-50% range. An attacker can send messages to any of these

users very easily. Even if the minor (registered as an adult) doesn’t specify — anywhere in Facebook — her high school and graduation year, and doesn’t make her friend list public to anyone, the attacker can still infer this information (using the techniques in this paper) and send a very personalized message to her.

2) *Contacting outside of Facebook:* For registered minors, the “Message” link is not available. Nevertheless, the attacker still may be able to contact the minor by guessing the minor’s email address [21]. Moreover, for many students, the first name, last name and city in the profiles can be linked with Skype profile information in the Skype directory, thereby providing a means of calling and videoconferencing with the minor [22].

3) *Establishing Friendship:* As the attacker knows the targeted minor’s gender, high school name, graduation year, many of her friends from the same graduation year, and potentially more information, the attacker has substantial resources for launching an *active* social engineering attack. For example, the attacker can create a fake Facebook account, masquerading as a student in the target high school (e.g., indicate he is a student in the high school, “like” Facebook pages related to the high school, and so on). As the Add Friend link is by default enabled for *all* Facebook users, the attacker can send friend requests to any of the discovered minors in the target high school. Some of the recipients will likely accept the requests [23]. Furthermore, after establishing friendship with a few users in the target high school, it becomes easier to establish friendship with the remaining users, as the probability of a user accepting friend requests increases significantly when the requestor can see some mutual friends [23]. Additionally, the social engineering attacks for OSNs described in [24] [25] can also be used to establish friendship with the victim. These active social engineering attacks, building on the passive high-school attack described in this paper, can be made to be fully automated. Once the attacker establishes friendship with the victim, not only can he send messages to the victim, but he can also extend the victim’s profile, since the information made available to friends is usually significantly richer than the information in a public profile.

VIII. WHAT IF COPPA NEVER EXISTED?

We now estimate how much privacy leakage there would be for minors in a world without the COPPA law. Without COPPA, there would be no age restrictions so that children under 13 could create accounts without having to lie about their ages. Although some children might still lie, as a joke or perhaps to obtain adult privacy privileges, the frequency of such occurrences would likely be much smaller. In this without-COPPA analysis, we will assume all users register with their actual birth dates. We also assume the OSN maintains the same privacy policy for minors as it does today — in particular, (*i*) when searching for users who attended

TABLE VIII
EXTENDING THE PROFILE FOR MINORS REGISTERED AS ADULTS

	HS1	HS2	HS3
# minors registered as adults	112	700	795
entire friend list public	73%	77%	87%
avg # of friends for users who make friend list public	405	960	908
public search enabled	71%	80%	86%
Message link	89%	86%	91%
relationship info	15%	26%	34%
interested in	13%	20%	33%
birthday	9%	4%	6%
average # of photos shared	19	51	57

a target high school or live in a target city, the OSN does not return registered minors; (ii) on a minor’s public profile page, the OSN displays only minimal information.

We now address two questions in the context of a world without COPPA. First, for a given target high school, can an attacker still find a set of OSN users such that (a) most of the students in the set attend the target high school (low false positive rate), and (b) the list contains most of students attending the high school (high coverage)? Second, for the students in the set, can an attacker create profiles that go significantly beyond the minimal public profiles? As we have seen, in a world with the COPPA law, the answer to both of these questions is yes. But to what extent is it also true in a world without COPPA?

In the high-school profile attack (with COPPA), a key component is discovering a set of “core” users who currently attend the target high school and have friend lists. Finding such a core set is facilitated by the fact that some minors have adult privacy privileges since they lied about their ages when creating accounts. Because Facebook treats these minors as adults, not only are they easier to find, but their friend lists are often public, thereby also making their classmates easier to find. *However, in our modified world without COPPA, such core users become more difficult to locate, since no minors would appear in search results.*

A. A Natural Attack in a COPPA-less World

Nevertheless, even in a world where everyone registers with their actual birth date, it would still be possible to locate some candidate minors. Here we discuss one natural mechanism to do this. Because many young adults (18-20 years old) will likely have friends who are a few years younger than them, if an attacker can find the young adults who recently graduated from the target high school and collect their friends, the attacker could create a list that contains minors in the target high school. Specifically, we suppose the attacker takes the following natural approach:

- 1) Obtain a set of users who are adults and have recently graduated from the target high school (or are adults in the last year of high school). Call the subset of these users who make their friend lists publicly available the core users.

- 2) Obtain the public profiles of all the friends of all the core users. Call the union of all these friends the candidate set. Most likely, the candidate set would contain many minors in the target high school.
- 3) All the minors (and some adults) will have minimal public profiles. To narrow down the candidate set, filter out all users who do not have minimal public profiles.
- 4) To further narrow down the candidate set, additionally filter out all users who have fewer than n friends in the core set. The attacker then considers this filtered set H as the minors in the target high school.

B. Apples-to-Apples Comparison

We now evaluate how successful this attack would be for a world without COPPA, and compare the success rates with those obtained in Section V for the world with COPPA. One of the challenges in this evaluation is that we are not able to collect data for the world without COPPA. If we apply the above heuristic to our existing data, then the above heuristic will not find any of the minors registered as adults; however, if we were applying the heuristic to actual without-COPPA data, more minors would be found since more minors would have minimal public profiles. In order to make a mostly apples-to-apples comparison, we therefore compare the number of minimal profile students obtained by the above heuristic for the without-COPPA case with the number of minimal profile students obtained in Section V for the with-COPPA case. We make this comparison for HS1, for which we have ground truth information. Recall HS1 has 325 students with Facebook accounts. Of these 325 students, 148 have minimal public profiles (22, 47, 45, and 34 students in 2012, 2013, 2014, and 2015, respectively). We now investigate how many of these 148 students are discovered in the two cases.

First consider the without-COPPA case. We will use students who have graduated in 2010 and 2011 to discover students graduating in 2012-2015, using the same data we used in Section V, which was collected in March 2012. Specifically, using the HS1 data obtained from the high-school search, we find 52 users who indicate their graduation year as either 2010 or 2011 from HS1 and publicly make available their friend lists. These 52 users become our “core set”. We then apply the heuristic above. For $n = 1$, we have

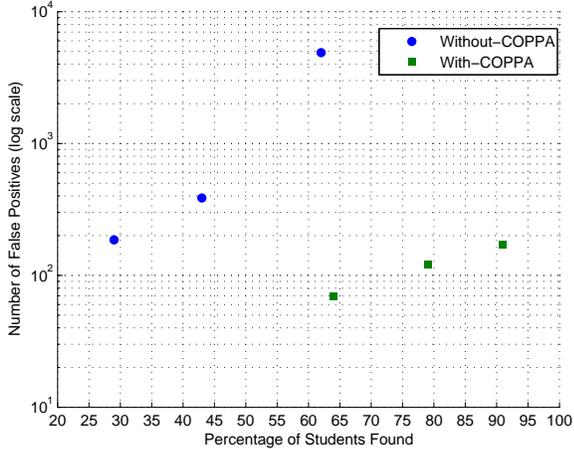


Fig. 3. With-COPPA vs. Without-COPPA

$|H| = 4,572$ filtered candidates and among these candidates, 92 of them are in the ground truth data set. Thus, $n = 1$ gives 62% of the minimal-profile ground-truth students with 4,480 false positives. The results for $n = 1, 2, 3$ are shown in Figure 3.

Now consider the with-COPPA case. For any t value, we need to determine the number of minimal profile students found and the number of false positives. Let M_t be the set of top- t users from Section V who have minimal profiles. The number of minimal profile students, z_t , is the number of students from M_t who are in the ground-truth set. The number of false positives is given by $|M_t| - z_t$. For $t = 300$, 165 users have minimal profiles of which 95 are in the ground truth data set. Thus, $t = 300$ gives 64% of the minimal-profile ground-truth students with 70 false positives. The results for $t = 300, 400, 500$ are shown in 3.

As shown in Figure 3, we see that without-COPPA, for obtaining the same number of minors as with-COPPA, the attacker has *many more false positives*. For example, with-COPPA gives 64% of the HS1 students with 70 false positives; without-COPPA gives 62% of the HS1 students with 4,480 false positives. Similarly, for the same number of false-positives, without-COPPA *finds significantly fewer students*. For example, with-COPPA gives 91% of the HS1 students with 170 false positives; without-COPPA only gives 29% of the HS1 students with 186 false positives.

We remark that *the differences between with-COPPA and without-COPPA would be even more pronounced for younger children in middle schools*. Our preliminary research shows that a variation of the high-school profile attack can be applied to many target middle schools (age 11-14). The attack would be much less successful in a world without COPPA, since adults 18-20 who have attended nearby high schools have relatively few friends in the target middle schools.

C. Profile Creation in a COPPA-less World

Having shown that, without COPPA, the attacker finds significantly fewer students for the same number of false positives, we now address the second question, namely, for the students in the attackers guess set H , can the attacker create profiles that go significantly beyond the minimal public profiles that Facebook displays for registered minors? Recall that the minimal profile at most contains name, profile picture, gender, and networks. In the without-COPPA case, using the heuristic above, the attacker would be able to augment this minimal profile with high school using the target high school, although his level of confidence would be significantly less because of the large number of false positives. Moreover, the attacker would not be able to easily determine the student’s graduation year, and the attacker would not be able to create a friend list that includes students in the same year, as is the case for with-COPPA. Thus, in the without-COPPA case, using the techniques in this paper, the attacker would not be able to construct a profile beyond the minimal profile plus (a low-confidence guess of) the high school. In the with-COPPA case, an attacker can obtain additional profile information for all minors: specifically, graduation year and high-school friend lists for registered minors; graduation year, high-school friend lists and often much more information (e.g., complete friend lists and shared photos) for minors registered as adults. Furthermore, for the without-COPPA case, an attacker would *not* be able to send Facebook messages to any of the minors (unlike the with-COPPA case).

In summary, under the assumptions in this section, and based on the heuristics in this paper, we can conclude that an indirect result of COPPA is that an attacker not only can discover more minors, but can also build much more extensive profiles than what would be the case in a world without COPPA.

IX. COUNTERMEASURES

In an ideal world, policymakers would enact laws and OSNs would take measures so that (i) it would be difficult for third parties to discover minors in targeted geographical regions and construct detailed profiles of those minors; while at the same time, (ii) providing a highly usable service for minors and adults alike. Designing and evaluating all combinations of possible laws and measures is a major research problem on its own. In this paper we examine just one promising countermeasure — namely, disabling reverse lookup — and quantify the reduction in privacy leakage.

With reverse lookup disabled, if a users friend list is hidden from strangers (either because the user has configured his friend list as such or because the user is a registered minor), then that user would not be visible to strangers in any other users friend list. If the OSN takes this countermeasure, users with hidden friendlist will not be found using reverse lookup, thereby reducing privacy leakages. To evaluate the

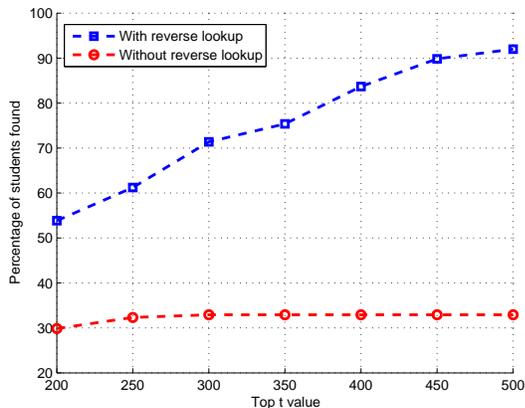


Fig. 4. Percentage of HS1 students found with and without reverse lookup

effectiveness of this measure, we estimate its impact on HS1, for which we have ground-truth information. Out of 325 ground-truth users of HS1, 112 of them make their friend list publicly available. For the remaining users, the friend list is not publicly available to strangers, and these users, by assumption, cannot be found via reverse lookup. So in evaluating this countermeasure, we remove these users from the candidate set and then proceed with the attack steps.

Figure 4 shows the results of the enhanced attack with filtering for HS1 with and without reverse lookup. From the results, it is clear that performance of high-school profiling attack decreases significantly. For example, by removing reverse lookup, the percentage of students found in the top-500 decreases from the 92% to 33%.

X. RELATED WORK

There are several earlier studies on the usage of OSNs by minors, both for minors (under 18) and underage users (under 13). In 2010, Pew Research released a report stating that 73% of online American teens ages 12 to 17 used an OSN website [26]. In 2011, Pew released another report, with collaborators at Cable in the Classroom and the Family Online Safety Institute, where it was found that 44% of online teens admit to having lied about their age so they could access a Web site or sign up for an online account [7]. Similar results have been reported for European teens [27]. Boyd et al. [9] provided survey data showing that many parents know their underage children are on Facebook in violation of the site’s restrictions, and that they are often complicit in helping their children join the site. More recently, Pew Research found that teens and adults have no significant variations for their privacy settings [28]. These reports collectively provide great insight into behavioral characteristics of minors and their parents. None of these reports, however, address automated discovery and profiling attacks on minors. Our work shows that because minors often lie to circumvent the age restriction, they put themselves and their non-lying high-school friends at risk for a variety of potential online and in-person abuses.

There is substantial previous work on using statistical inference to infer private information about OSN users. Zheleva and Getoor [29] proposed techniques to predict gender and political views of users in four real-world datasets (including Facebook) using general relational classification and group-based classification. Jernigan and Mistree [30] demonstrated a method for accurately predicting the sexual orientation of Facebook users by analyzing friendship associations. Other papers [31], [32], [33] have also examined inferring private information from social networks. Thomas *et al* examine scenarios where conflicting privacy settings between friends will reveal information that at least one user intending to remain private [34]. Becker and Chen [35] inferred many different attributes of Facebook users, including affiliation, age, country, degree of education, employer, high school name and grad year, political view, relationship status, university and zip code using the most popular attribute values of the user’s friends. Dey et. al [20] examine a large dataset and develop a methodology to estimate ages of Facebook users. Mislove et al. [36] proposed a method of inferring user attributes by detecting communities in social networks, based on the observation that users with common attributes form dense communities.

All of the above studies focus on inferring information about adults. To our knowledge, this is the first paper that identifies the privacy problem in OSNs for minors, and also the first paper to quantify the extent of the privacy leakage. The problem is challenging since, for registered minors, little information, including friend lists, is available to an attacker. The attack makes use of two key properties in modern OSNs: (i) many minors lie about their age and are therefore considered adults by the OSN; and (ii) using reverse lookup, an attacker can construct a user’s friend list even if the user hides her friend list to everyone. This is also the first paper to measure the additional privacy disclosure risk for minors due to the enactment of the COPPA law [37].

There is also a body of work related to using social networks to perform social engineering attacks on users. Jagatic *et al* [38] collected friendship links by crawling OSNs, sent spoofed email messages from one friend to another, and redirected the recipient to a phishing site. They showed the recipients are four times as likely to become victims if they are solicited by someone appearing to be a known acquaintance. Bilge *et al* [24] investigate cloning an existing user profile and sending friend requests to the clone’s friends. They showed that users are more likely to accept friend requests when coming from someone they believe they know. Irani *et al* [25] studied Reverse Social Engineering (RSE) attacks in OSNs, whereby the attacker creates a persona that should be attractive to the victim, so that the victim initiates a friendship request with the persona. The authors showed much higher success rates with RSE are possible then when actively contacting the victim. These social engineering attacks are complementary to the

passive attacks in this paper. In particular, an attacker can first begin with the high-school profiling attack, obtaining an extensive list of target students with varying degrees of profile information. The attacker can then use the student list and profiles for phishing, cloning, establishing online friendships, and setting the stage for RSE attacks.

XI. CONCLUSION

In this paper we have shown how a privacy law for protecting children's privacy can inadvertently increase minor's exposure to third parties. Facebook and other Online Social Networks (OSNs) take precautions to prevent strangers from using their services to extensively profile minors. But because a significant fraction of minors lie about their ages, we show how many of the precautions can be circumvented, putting both lying and truthful minors at risk. For a given target high school, we described an attack of using an OSN to profile the current students in the high school. The attack finds the majority of the students in the school, and for each student builds a profile that includes information that is not normally available to strangers, including current school, graduation year, and high-school friends. The profiles of the identified minors also include a varying amount of additional information, including shared photos and wall postings.

We estimated how much privacy leakage would occur in a world without the COPPA law and compared the estimate to the extent of leakage in our current world with COPPA. Our results suggest that an attacker not only can discover more minors, but can also build more extensive profiles than what would be the case in a world without COPPA. Thus, in terms of third-party privacy attacks, COPPA actually puts minors at greater risk than they would be if the law had never been enacted.

Although the COPPA law indirectly exacerbates the third party privacy problem for minors, we are certainly not arguing that governments should abandon enacting laws to protect the online privacy of children. We believe, however, that the laws must be carefully designed and consider leakages to third-parties as well as to first-parties.

REFERENCES

- [1] "Children's Online Privacy Protection Act," 1998, <http://www.ftc.gov/ogc/coppa1.htm>.
- [2] "Groups Make Recommendations for Kids' Facebook," Adweek, June 18, 2012, <http://www.adweek.com/news/technology/groups-make-recommendations-kids-facebook-141195>.
- [3] "Do Not Track Kids Act of 2011," May 13, 2011, <http://www.gpo.gov/fdsys/pkg/BILLS-112hr1895ih/pdf/BILLS-112hr1895ih.pdf>.
- [4] "Update Urged on Childrens Online Privacy," The New York Times, September 15, 2011, <http://www.nytimes.com/2011/09/16/technology/ftc-proposes-updates-to-law-on-childrens-online-privacy.html>.
- [5] "Silicon Valley Objects to Online Privacy Rule Proposals for Children," The New York Times, November 5, 2012, <http://www.nytimes.com/2012/11/06/technology/silicon-valley-objects-to-online-privacy-rule-proposals-for-children.html>.
- [6] "How does privacy work for minors?" <http://www.facebook.com/help/?page=214189648617074>.
- [7] A. Lenhart, M. Madden, A. Smith, K. Purcell, K. Zickuhr, and L. Rainie, "Teens, kindness and cruelty on social network sites," Pew Internet, November 9, 2011, http://pewinternet.org/~media/Files/Reports/2011/PIP_Teens_Kindness_Cruelty_SNS_Report_Nov_2011_FINAL_110711.pdf.
- [8] "Consumer Reports survey: 7.5 million Facebook users are under the age of 13, violating the sites terms," Consumer Reports, May 10, 2011, <http://pressroom.consumerreports.org/pressroom/2011/05/cr-survey-75-million-facebook-users-are-under-the-age-of-13-violating-the-sites-terms-.html>.
- [9] danah boyd, E. Hargittai, J. Schultz, and J. Palfrey, "Why Parents Help Their Children Lie to Facebook: Unintended Consequences of the Childrens Online Privacy Protection Act," *First Monday*, vol. 16, no. 11, November 7, 2011.
- [10] "On the Web, Children Face Intensive Tracking," The Wall Street Journal, September 17, 2010, <http://online.wsj.com/article/SB10001424052748703904304575497903523187146.html>.
- [11] "Senator Opens Investigation of Data Brokers," The New York Times, October 10, 2012, <http://www.nytimes.com/2012/10/11/technology/senator-opens-investigation-of-data-brokers.html>.
- [12] "Attorney General Kelly announces criminal charges in elaborate 'Facebook' false identity scam targeting young girls for sex," February 10, 2012, <http://www.attorneygeneral.gov/press.aspx?id=6431>.
- [13] "Find Friends Portal," <https://www.facebook.com/find-friends/browser/>.
- [14] "Age requirements on Google Accounts," <http://support.google.com/accounts/bin/answer.py?hl=en&answer=1350409>.
- [15] "Creating a Google Plus Account Now Requires You to Enter Your Birthday," August 27, 2011, <http://techie-buzz.com/social-networking/google-age-restrictions.html>.
- [16] "Google+ Search by School Portal," <https://plus.google.com/circles/school>.
- [17] "Google+ Teen Safety Guide - Features for teens," <http://support.google.com/plus/bin/answer.py?hl=en&answer=2409072>.
- [18] P. Sen, G. Namata, M. Bilgic, L. Getoor, B. Gallagher, and T. Eliass-Rad, "Collective Classification in Network Data," *AI Magazine*, vol. 29, no. 3, pp. 93–106, 2008.
- [19] S. Chakrabarti, B. Dom, and P. Indyk, "Enhanced hypertext categorization using hyperlinks," in *Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data*, 1998, pp. 307–318.
- [20] R. Dey, C. Tang, K. W. Ross, and N. Saxena, "Estimating age privacy leakage in online social networks," in *Proceedings of the IEEE INFOCOM 2012, Orlando, FL, USA*, 2012, pp. 2836–2840.
- [21] M. Balduzzi, C. Platzer, T. Holz, E. Kirda, D. Balzarotti, and C. Kruegel, "Abusing social networks for automated user profiling," in *Proceedings of the 13th International Conference on Recent Advances in Intrusion Detection*, 2010, pp. 422–441.
- [22] S. Le Blond, C. Zhang, A. Legout, K. Ross, and W. Dabbous, "I Know Where You are and What You are Sharing: Exploiting P2P Communications to Invade Users' Privacy," in *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, 2011, pp. 45–60.
- [23] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "The social-bot network: when bots socialize for fame and money," in *Proceedings of the 27th Annual Computer Security Applications Conference*, 2011, pp. 93–102.
- [24] L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda, "All your contacts are belong to us: automated identity theft attacks on social networks," in *Proceedings of the 18th International Conference on World Wide Web*, 2009, pp. 551–560.
- [25] D. Irani, M. Balduzzi, D. Balzarotti, E. Kirda, and C. Pu, "Reverse social engineering attacks in online social networks," in *Proceedings of the 8th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, 2011, pp. 55–74.
- [26] A. Lenhart, K. Purcell, A. Smith, and K. Zickuhr, "Social media and mobile Internet use among teens and young adults," Pew Internet, February 3, 2010, http://www.pewinternet.org/~media/Files/Reports/2010/PIP_Social_Media_and_Young_Adults_Report_Final_with_toplevels.pdf.
- [27] S. Livingstone, K. Olafsson, and E. Stakrud, "Social Networking,

- Age and Privacy,” EU Kids Online, 2010, <http://www2.lse.ac.uk/media@lse/research/EUKidsOnline/ShortSNS.pdf>.
- [28] M. Madden, “Privacy management on social media sites,” Pew Internet, February 24, 2012, http://www.pewinternet.org/~media/Files/Reports/2012/PIP_Privacy_management_on_social_media_sites_022412.pdf.
 - [29] E. Zheleva and L. Getoor, “To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles,” in *Proceedings of the 18th International Conference on World Wide Web*, 2009, pp. 531–540.
 - [30] C. Jernigan and B. F. T. Mistree, “Gaydar: Facebook Friendships Expose Sexual Orientation,” *First Monday*, vol. 14, no. 10, 2009.
 - [31] W. Xu, X. Zhou, and L. Li, “Inferring Privacy Information via Social Relations,” in *24th International Conference on Data Engineering Workshop*, 2008, pp. 154–165.
 - [32] J. He, W. W. Chu, and Z. V. Liu, “Inferring privacy information from social networks,” in *Proceedings of the 4th IEEE International Conference on Intelligence and Security Informatics*, 2006, pp. 154–165.
 - [33] C. Tang, K. Ross, N. Saxena, and R. Chen, “What’s in a name: a study of names, gender inference, and gender behavior in Facebook,” in *Proceedings of the 16th International Conference on Database Systems for Advanced Applications*, 2011, pp. 344–356.
 - [34] K. Thomas, C. Grier, and D. M. Nicol, “Unfriendly: multi-party privacy risks in social networks,” in *Proceedings of the 10th International Conference on Privacy enhancing technologies*, 2010, pp. 236–252.
 - [35] J. Becker and H. Chen, “Measuring Privacy Risk in Online Social Networks,” in *Proceedings of W2SP 2009: Web 2.0 Security and Privacy*, 2009.
 - [36] A. Mislove, B. Viswanath, K. P. Gummadi, and P. Druschel, “You are who you know: inferring user profiles in online social networks,” in *Proceedings of the third ACM International Conference on Web Search and Data Mining*, 2010, pp. 251–260.
 - [37] Anonymous, “The High-School Profiling Attack: How Online Privacy Laws Can Actually Increase Minors Risk,” Tech. Rep., November, 2012.
 - [38] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, “Social phishing,” *Commun. ACM*, vol. 50, no. 10, pp. 94–100, Oct. 2007.