Fakesbook
A social networking platform for teaching security and privacy concepts to secondary school students

Maximilian Zinkus, Oliver Curry, Marina Moore, Zachary Peterson, Zoë J. Wood
Computer Science - California Polytechnic State University
San Luis, Obispo, CA
mzinkus,owcurry,mmoore32,znjp,zwood@calpoly.edu

ABSTRACT
As frequent users of social networking applications, middle and high school students are well-suited for curricular interventions that leverage these technologies. Allowing students to see “behind the curtain” of these applications provides them with a unique opportunity to better understand the discipline of computer science upon which these technologies are built and influences their perceptions of computer security and privacy. We present a novel social networking simulation that allows students to create a social network account, including profile data and images, and to manage privacy settings and friend connections. The platform, named Fakesbook, presents students with a visualization of the social network as a graph, enabling them to observe the spread of profile data (theirs and others’) depending on friend connections and choices of privacy settings. We additionally present our lab curriculum which uses Fakesbook to enable active learning and adversarial thinking to engage students and build agency with regard to privacy and computing concepts. We deployed and, over several years, evaluated our platform and curriculum with hundreds of students from a diverse set of backgrounds at educational events designed to introduce these populations to computer science, cybersecurity, and privacy. Survey results indicate that students gained or deepened their understanding of online privacy and security and that 86% of participants found that Fakesbook helped them “think about privacy and computer security.”

KEYWORDS
security and privacy, computer science education, privacy and social networking applications

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACM Reference Format:

1 INTRODUCTION
Social networking has become ubiquitous for people of many backgrounds and ages, including a large percentage of minors and young adults. Facebook alone has self-reported over two billion monthly active users [3], and recent literature estimates around 75% of all children ages 12-17 interact with some form of social networking regularly [7, 14]. Concerningly, teenagers with fully public profiles are more likely to experience online cruelty or harassment [7].

Our work stems from the notion that by fostering understanding of online privacy and risk management we can encourage students to adopt safer privacy practices. Additionally, as computer science educators, we aim to build a platform that allows students to see “behind the curtain” to enable deeper understanding of social networking applications. We do this in two ways. First, we provide a platform that visualizes a graph with all users, within a small classroom setting, as nodes, and “friend” relationships as edges. Using the graph visualization, we enable students to see how privacy settings and social connections translate into their data being accessible to a wider and wider group of people. Second, we create a lab experience that highlights the power of programming and contextually introduces computer science as a discipline. Specifically, we want students to understand that data privacy and access control can be controlled programmatically, and thus we include an introductory encryption exercise using the Python programming language. This exercise provides a context-based (privacy and security) introduction to basic programming concepts (variables, loops, arrays and arithmetic operators).

We present the development of our social networking application, Fakesbook, our curriculum, and evaluation through instructional use of the platform over multiple years and with hundreds of students. Fakesbook was used in one-time lab experiences of up to two hours with diverse groups of 20-25 secondary school students of various ages. For each lab experience, we ran Fakesbook on a local network to avoid sharing student information on a public system, while providing students freedom to explore the consequences of various privacy settings. These lab experiences, guided by our curriculum of exercises on the Fakesbook platform and introductory encryption programming lessons, allow secondary school students to gain insight into social networking. Post-lab surveys indicate that the activities and platform were successful in enhancing students’ understanding and value of privacy and computer security. Our platform is made available via GitHub for interested educators at https://github.com/Fakesbook/Fakesbook.

2 RELATED WORK
Content posted on social networking sites can have real consequences, evidenced by many who report regretting posts they have
made [21]. However, users of social networks often have little understanding of the connection between privacy settings and their data, and education about the differences even at the college level is lacking [8, 10]. Using our platform and a pedagogical approach which draws on strategies of visualization and adversarial analysis, we seek to encourage safer privacy practices in social network users at an early stage. Visualization, when paired with active learning [5] and adversarial analysis [11], has been shown to benefit computer science and cybersecurity education. Specific examples include cryptography [13, 17], network security [1, 22], phishing [15], password security [23], and software security [12]. Some of these visualizations are highly detailed, at the expense of accessibility. Tao et. al. [17] and Schweitzer [12] provide visualizations in cryptography and software security respectively, both of which are information-dense. However, they are likely inaccessible to our target audience. Atwater et. al. [1] provide visualization in network security, using a graph-like format to describe a communication network. We employ a similar visual pattern to describe a social network. Zhang-Kennedy et. al. [23] employ a generated interactive narrative along with visualization to convey password security and found that the narrative was helpful for impact and memorability. Our associated lab curriculum overlays a narrative on the Fakesbook application with the goal of similarly improving outcomes. Ours is the first work that we are aware of to use visualization for communicating privacy concepts.

Regarding privacy pedagogy in general, the U.S. National Security Agency and Department of Homeland Security (NSA/DHS) jointly provide certification for institutions of higher education of program content including privacy [20]. They provide associated Knowledge Units [19] which document baseline topic coverage. Our lab experience, consisting of the Fakesbook application and associated curriculum, addresses the first two of the four listed outcomes in the Privacy Knowledge Unit, and additional curricula can expand it to address the other two more fully. As these certifications are usually applied to tertiary educational programs, we consider this to be strong support for our curricular content.

3 OVERVIEW

Our aim with this work is to communicate the importance of privacy and security in social networking applications and to encourage students to engage with these concepts and disciplines. To provide students the opportunity to explore privacy settings in a social networking platform, we developed Fakesbook, a web application that mimics the basic behavior of a typical social networking application. However, unlike a typical social networking application, we visualize the network graph and highlight the sets of users who can see each element of a user’s profile data. This software can be used to introduce privacy concepts to a general audience, and to encourage them to make proactive decisions on what to share on these kinds of platforms. We leverage our platform through our associated curriculum comprised of lab activities.

4 THE FAKESBOOK PLATFORM

Platform overview: Fakesbook is a simplified version of commercial social networking applications designed to run privately, on a local network, and intended only for educational use. Our platform allows students to create mock social networking accounts, including basic information about themselves and a profile image. After a brief registration and profile data population process, students are directed to a single view containing the remaining features of the platform. This view displays the user population graph with ‘friend’ connectivity, a profile information panel, and a privacy settings panel. An example of the complete view is shown in Figure 3.

Steps for platform use: The Fakesbook platform is intended to be used in a directed lab setting. In our curriculum, students are guided through the following steps:

- **Create a user account (profile):** Supported profile data fields include name, favorite color, age, gender, interests, hometown, and an image. The data population screen step compromises its own screen in the platform. In addition, users may edit their profile data fields when viewed/selected in the main screen.
- **Create the social network (‘friending’):** Once a user has created an account, their main view is a large screen that shows all other users who have created an account on Fakesbook. The platform allows for ‘friending,’ where students can connect with their peers by selecting a target user in the network view and clicking ‘Add Friend.’ The request is pending until the target user accepts, emulating popular social networks and somewhat mitigating excessive friending which may distract from the exercise.
- **Modify privacy settings:** Students can control the visibility of their data by editing the relevant privacy settings for their
profiles, as shown in Figure 2. Privacy settings determine if the data is visible to everyone, friends of friends, or just friends.

Social Network Graph Visualization: Visualizations are designed to highlight graph connectivity and the effect of privacy settings on the propagation of profile data. Students can select (click on) any node in the graph to view that user’s profile (as determined by that user’s privacy settings). Figure 1 shows an example of viewing another user’s data before and after “friending”. Students can also drag nodes around to move the graph, which allows them to better see connections, as highly-connected graphs can lead to obscuring overlap. D3.js and additional JavaScript is used to give the visualization basic physical interactivity, through features such as collision resistance of nodes and elasticity of edges.

Profile Data Visualization: When students edit their data privacy settings, they can see exactly who else has access to view each element of their profile. This is a core component of Fakesbook as it allows students to see how easily their data can be made visible to others and especially the result of permissive default privacy settings (i.e. everyone). When a user interacts (clicks or hovers their mouse) with their account privacy settings, Fakesbook displays a visualization that highlights the users with access to the profile item related to that privacy setting. Color coding (green, yellow, orange) is used to help key the user into the meaning of having their data more widely available. In this regard, we quietly exploit common associations of green with “good” and orange with “warning.” We justify these associations upon the observation that teenagers with fully public profiles are more likely to experience online cruelty or harassment[7]. However, to mitigate concern, we don’t include this in the student discussion. Figure 3 illustrates sets of users in the social network graph who can see various elements of a profile with each privacy setting option. The top image in Figure 3 is set to friends only, whereas the middle image is set to friends of friends, and the bottom image is set to everyone. In addition, when a user interacts with their own profile data, the same graph visualization appears to strongly associate profile data with other user’s access. These visualizations and immediate visual feedback when changing privacy settings encourages students to realize their agency in controlling data access to others.

User Profile Data Storage: During the use of the Fakesbook platform, an SQLite [16] database is used to store profile data. This data can be exported as an ASCII text file as shown in Figure 4. This text data can then be shown to students and used in additional curricular exercises as described in Section 5. The instructor can export the current database at any time and load new databases onto the platform. Additionally, Fakesbook includes a secondary tool which can be used to populate the database with a network of generated users, which is intended to be used if example user profiles or graphs with specific shape are required.

Figure 2: Fakesbook privacy settings panel.

Figure 3: Top: example Fakesbook graph with the friends of the currently logged in user highlighted in green. Middle: example Fakesbook graph with the friends of friends and friends of the currently logged in user highlighted in yellow. Bottom: example Fakesbook graph with everyone highlighted in orange (profile image from public domain).

Figure 4: An example of Fakesbook user data saved and stored as an ASCII text file. Such a file can be used to start a lab without the need for user registration and can also be shown to students and used in subsequent exercises.

Implementation: Fakesbook is backed by a Python/Flask [4] webserver, uses the Twisted [18] module for parallelization, an SQLite [16] relational database, and an extensive D3.js [2] frontend. The Flask webserver communicates with the database to store and update user data, and to serve that user data upon request over HTTP. The Twisted layer wraps the Flask webserver in order to parallelize the application through multi-threading to maintain responsiveness under load. The D3.js frontend runs in the user’s browser, rendering the user interface and visualizations while requesting the
latest social network graph data from the server, and forwarding user inputs such as registration to Flask for storage. See Figure 5 for an overview of the platform architecture.

Figure 5: Fakesbook platform architecture.

This lightweight client-side rendering architecture, paired with relatively small class sizes (~25 students), allows us to instantiate the platform in constrained environments, such as on a laptop in a classroom, while providing near real-time updates to graph structure visualizations. The user interface provides immediate feedback for user actions where possible by allowing the visualization to update asynchronously with the backend.

Keeping to information security principles, privacy settings are evaluated in Flask rather than in the frontend, and the data sent to each client is replaced with Hidden where appropriate.

5 PEDAGOGICAL APPROACH

Instructional setting: Fakesbook is meant to be used in a lab setting under direction from an instructor, and our curriculum is intended for small classes with access to workstations with web browsers that can be directed to a local server running the platform. Instructor presence is critical in that reflection and active learning activities are central to our curriculum, and these participation-based activities are challenging when class sizes grow larger. However, the platform is not restricted to a particular curriculum and we hope can be extended to support additional learning outcomes.

Warm-up: To motivate students to start thinking about social networks and how people and data might be connected, the lab starts by asking students to use pen and paper to draw out their own immediate social network. Figure 6 shows an example network created by a participating student. This warm-up guides students to start thinking about graphs, social networks, and privacy in a way that is inclusive and scaffolded, supporting students with perhaps little interest or familiarity with the technologies that implement social networking. We then transition to a discussion about social networking companies and applications to set our context.

Building a social network (organic): Next, we direct students to create a profile on Fakesbook, leading them through profile creation and encouraging realistic inputs for "interests," while allowing them to withhold other information they may feel uncomfortable sharing. Once accounts are made, we allow for the evolution of an organic graph structure as students ‘add friends’ with their in-class peers and discover Fakesbook features. An example of an organic graph is shown in Figure 7. We ask students to observe graph structure and discuss how they went about creating connections. Then we encourage the students to check who else can access their profile data, and change their privacy settings and observe the effects of the changes. Through this interaction we leverage the studied benefits of visualization paired with active learning to impart familiarity with and encourage critical analysis of privacy settings [5, 9]. Reflection questions at this stage include:

- How connected is the graph?
- Are any friend-of-a-friend connections surprising?

Building a social network (guided): In our experience, we have found that the majority of the student groups tend to create atypical social networking graphs (i.e. overly inclusive to all, with everyone friending everyone else, or overly reticent with few connections between those present). Essentially, creating a social network constrained to the fellow participants of the lab setting, which might include many unfamiliar peers, is abnormal to normal social networks with much larger sets of users. As we want our subsequent exercises to be applied to a more typical social network [6] (with clusters of highly connected groups and few edges between these groups), we then lead the students to create a new graph. As before, students create a ‘realistic’ profile, however, they are required to
list at least three interests in their profiles and to use the ‘hometown’ field to simulate ‘checking in’ to a location (i.e. Starbucks Cafe, school, etc., to mimic the ‘checking in’ feature of many social networking applications). Then we divide the students into small groups and task them with ‘friending’ all other members within their group. Once each group is fully connected we select one or two students from each group and ask them to friend another student in a different group, continuing in this manner until links between all the groups are created. Figure 8 shows an example of one of these graphs. Students are asked to review and update their privacy settings as they might typically set them, with the requirement that everyone have their profile image be visible to everyone.

Finally, we ask students to perform an adversarial analysis of profile data use. Students are asked to use the platform to discover useful marketing or identifying information about students in the other groups. For example, students are asked to identify where each member of other group is located and potential products to advertise to the other group members. Students usually exhibit a high level of engagement as they figure out how to exploit the graph and connections to learn about one another. Participants commonly express surprise at how easy it was to find information about one another, either directly or via friends (and friends of friends). We conclude this lesson by discussing results and strategies and revisiting social networking applications motivations with regards to data collection and security.

Context-based introduction to programming: To further help students see behind the curtain, we show students the text file that demonstrates how profile data can be digitally stored. Using this as a starting point, we transition to an exercise focused on data and encryption. We lead students through an introductory programming exercise to encrypt and decrypt this text data using Python and a Caesar cipher. We provide and carefully discuss base code to introduce variables, loops, characters, arrays, and the necessary mathematical operators. This final lab exercise is meant to empower students to see that computer programming is accessible and that they can control how data might be stored and communicated between computers. At this point, we also show students a key segment of code from the Fakesbook platform, shown in Figure 9.

This code segment is associated with privacy control and sharing it allows us to emphasize the connection between the platform itself, programming, and privacy and security concepts.

6 RESULTS

Fakesbook has been used over a two-year period to expose hundreds of secondary school students to concepts related to cybersecurity and privacy. In our experience, the platform generates a good deal of student excitement, engagement, and ultimately learning for the lab participants.

Fakesbook was first used in Spring 2017 during an event designed to introduce secondary school students to computer science. Four sections of ∼25 students used the platform in its alpha state for a short lab of 30 minutes and provided feedback.

Fakesbook was then used in Summer 2017 in the ‘cyber track’ for three separate week-long summer camps targeted at introducing secondary school students to engineering. Specifically, the lab was used in Cal Poly, EPIC: Engineering Possibilities in College, which is a unique summer program for middle- and high-school students. The camp aims to include 50% underrepresented groups in engineering including female, first-generation, and low-income students. Each Fakesbook lab was a two-hour experience using our curriculum with three groups of ∼20 students participating. Anecdotally, students were positively engaged in the exercises, and particularly enjoyed the adversarial mining of other students’ data. Based on observations, further revisions to the lab were made, including adding more instructor direction on friending to produce more typical social network graphs. A desired attribute is fewer friend connections per node, as excessive connectivity blurs the distinction between the friends of friends and everyone groups and is atypical in real social network graphs [6].

Fakesbook was used again in Summer 2018 in both a designated ‘cyber’ track and a general ‘engineering’ track over a three-week period for EPIC. Six groups of ∼20 students participated in the lab. Students completed pre- and post-lab surveys assessing their attitudes on online privacy and security and their thoughts on the Fakesbook platform. Pre-lab questions included:

(1) “Have you ever changed the privacy settings from the default on any of your social media accounts?” (yes: 91.7% No: 6.6%)

Figure 8: A Fakesbook graph showing an example of a graph generated by separate groups of students.

Figure 9: Fakesbook code shown to students relating programming concepts to the platform.

for k in perms:
  if perms[k] == 0:
    if is_friend:
      show[k] = usermap[k]
  else:
    show[k] = 'hidden'
  elif perms[k] == 1:
    if is_friend or is_fof:
      show[k] = usermap[k]
  else:
    show[k] = 'hidden'
  elif perms[k] == 2:
    show[k] = usermap[k]
Post-lab questions included:

1. "Overall, how important is it to you to keep your online data private? (images, status updates, location)?" (5-point Likert scale; ≥4: 81.6%, 3: 2.4%, 4: 30.6%, 5: 50.4%)

2. "How likely is it to talk to your parents or your friends about data and privacy online?" (5-point Likert scale; ≥4: 50%, 3: 27.8%, 4: 27.8%, 5: 22.2%)

3. "Did working with the ‘fakesbook’ application and seeing a graph of a social network help you think about privacy and computer security?" (yes: 86.1%, maybe: 11.1% no: 2.8%)

4. "Any comments or final thoughts?" (free response, see below)

The pre-lab surveys indicate that students were already thinking about privacy and social media prior to the lab, while post-lab surveys indicate they deepened their knowledge and engagement from participating in the lab. When asked in the pre-lab survey, "Have you ever changed the privacy settings from the default on any of your social media accounts?", eight of the 121 students indicated that they had never changed their privacy settings, two indicated ‘N/A’ (likely students without social media accounts), while the remaining 111 students responded with ‘yes’. This indicates that participants started their lab experience with some engagement with social networking privacy settings.

When asked in the pre-lab survey, "Overall, how important is it to you to keep your online data private? (images, status updates, location)?" on a five-point scale, prior to the lab, the average score was 4.22 (with a standard deviation of 0.98). After the lab, students increased their self-evaluation of the importance of online privacy, responding to the same question with an average score of 4.57 (standard deviation of 0.72). This increase in the importance of online privacy indicates that the students gained respect for the importance of privacy and security due to the lab activities.

We also asked in the post-lab survey about the utility of the Fakesbook platform. Specifically, "Did working with the 'fakesbook' application and seeing a graph of a social network help you think about privacy and computer security?" Participants responded very positively, with 86% responding that it was helpful, and 11% responding that 'maybe' it was helpful.

In addition, the post-lab survey allowed for students to voluntarily add "Any comments or final thoughts." Out of the 55 voluntary written comments about the lab, 45 (80%) were very positive, along the lines of "this was my favorite lab" or "The 'fakesbook' application was extremely helpful in visualizing online connections!". The remaining 10 comments primarily related feedback about the 'friending' process in Fakesbook or the challenges of the Caesar cipher python coding portion of the lab, which we are taking into consideration for future work. The fact that close to 40% of all participants voluntarily wrote very positive comments attests to their enthusiasm for the lab activity.

Another way to measure the impact of the lab experience is to examine student’s plans for reflecting on the lab and sharing the information with others. In the post-lab survey we asked students about the likelihood of them talking to their parents and friends about privacy. The majority, 77.8%, of students rated that they were likely or very likely to speak to others about the topic, which indicates the positive impact of the lab experience for participants.

7 CONCLUSION & FUTURE WORK

Minors and young adults today are creating data-rich online presences on social networks. Through a combination of our platform and curriculum, we seek to convey familiarity with the concepts of online privacy and risk management and to promote student understanding of their agency in these regards. We devise our approach to leverage the well-studied pedagogical efficacy of interactive visualization and adversarial thinking. Through this combination, we seek to promote online safety for young users who might put their personal data at risk before fully understanding the implications of their actions. It is our goal that this exercise will empower students to control their data and the context-based programming assignment might encourage them to consider learning more programming. Fakesbook is available on GitHub, https://github.com/Fakesbook/Fakesbook and interested educators can contribute to the project via use and software development.

Future work includes supporting self-guided student exercises, e.g. "You just changed your privacy settings, how many people do you think can see your hometown now", which would enable the platform to support embedded assessments and more independent student exploration.

In addition, to validate our progress toward privacy education, future work includes following up with students after their experience with the platform. The majority of students indicated that they were likely to speak to others about the topic, however, the addition of follow-up interviews would allow us to better measure the impact of our tool on short- and long-term behavior.

Finally, we envision that this platform could be extended as an educational tool. For example, further work on the graph visualization could help students see the relationships between sub-groups in the graph. In addition, we are interested in extending the visualizations on the platform to communicate graph theory concepts which are central to common algorithms in computer science.

ACKNOWLEDGMENT

The authors would like to thank the organizers and volunteers of Expanding Engineering and of EPIC. This paper is based on research supported by the National Science Foundation under Grant No. 1628726. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES


