Teaching Students to Be Internet Stalkers - Experiences From An Open Source Intelligence Class Project

Mark Shaneck
mshaneck@liberty.edu

Gwen Shaneck
gshaneck@liberty.edu

Liberty University
1971 University Blvd.
Lynchburg, VA 24515

Abstract - Teaching Cyber Security students to think like an attacker and attempting to inculcate in them the security mindset is a common yet important facet of security education. Most of the literature concerning this topic focuses on technical activities such as network mapping, reconnaissance, or host exploitation. In this paper, we describe our experience with a hands-on activity in which students had the chance to think like an attacker, but in the less technical, but arguably equally important, area of Open Source Intelligence gathering. Our students were given the opportunity to conduct an information gathering attack on real targets, compiling the results into a report that was delivered to the target, and receiving feedback from the target about their findings. We finished the project by surveying the students to discover what they learned and how the project changed, not only their perspective on the availability of information on the internet, but also how they intend to change their behavior as a result of the exercise.

Categories and Subject Descriptors

K.3.2 [COMPUTERS AND EDUCATION]: Computer and Information Science Education.

General Terms

Open Source Intelligence Gathering, Privacy

Keywords

Open Source Intelligence, Data Privacy, OSINT, Offensive Information Security, Security Education, Privacy Education, Security, Education, University

1. INTRODUCTION

Security education is an important part of undergraduate Computer Science and Information Technology curricula. One of the most common recurring themes about incorporating security education into courses is the need to teach the "security mindset" [5, 2], in which the students are put into situations where they have the chance to "think like an attacker". One of the more interesting examples of this is where students are required to cheat on a test [3]. In many cases, researchers emphasize the importance of hands-on, practical activities [10, 7, 6]. One compelling perspective put forward by Gava, et al. [5] describes the need for challenge-based learning, highlighting Cyber Security competitions and challenges. In light of this research, we try to incorporate these sorts of activities into our courses as much as possible. However, to the best of our knowledge, with the exception of the cheating test [3], the primary topics that were covered in these hands-on activities were technical in nature [10]. In this paper, we describe an assignment we gave that provides a hands-on activity in an important, yet less technical, area of Cyber Security that appears to have been overlooked in the Cyber Security Education research.

Open Source Intelligence (OSINT) is an important step in any offensive information security engagement. Not only is it necessary to gather information about a target before attempting to penetrate their network, it is also a very effective source for information that can be used in targeted phishing attempts. Individuals need to be aware of the amount of information that is publicly available through open sources and even relatively simple Google searches. Many people, particularly college students, are unaware of the breadth of information that can be gleaned about them on the internet, as well as the potential damage that can be caused by seemingly unimportant pieces of data.

In the fall of 2015, we assigned an OSINT collection project for our Computer and Network Security course consisting of seniors pursuing Computer Science degrees. This project was designed to introduce students to the tools and techniques of OSINT, educate them on the prevalence of information that is publicly available, give them experience in performing this sort of assessment on a real individual, and to get feedback from their targets on the value of the information that they were able to glean. While our students were aware that there was a large amount of information that was accessible, our hypothesis was that they did not fully appreciate the extensive nature of the information, the sensitivity of some of it, nor how dangerous it can be.

The rest of this paper is organized as follows. Section 2 discusses the overall project and how it was structured, including what types of tools were covered in the lectures preceding the assignment. Section 3 presents the information that the students were able to find, the spear phishing attempts that they believed would be most successful, and the targets' feedback about the accuracy of their information. Section 4 discusses the learning outcomes of the project from the students' perspectives and their future behavioral changes based on what they have learned. Finally, Section 5 concludes the paper and identifies ways to improve this project in the future.

2. PROJECT OVERVIEW

The Computer and Network Security course in which we assigned this project had 24 students enrolled. These students were divided into four equal teams consisting of six students. Each student group was tasked with conducting OSINT research on a real person. There were four different targets, hereafter known as Target 1, Target 2, Target 3, and Target 4. Each target was involved in the security field in some capacity and were all security-minded individuals. None of the targets lived in the local area. Each target was informed of the project prior to the beginning of the semester and each agreed to serve as a target for the duration of the project.

The students were provided the name and a general location for their target, along with the information that the target was involved in the Cyber Security industry in some capacity. The student groups were instructed to find as much information about their target as they could using only open sources. They were free to subscribe to any online information service, at their own expense, although none of the student groups availed themselves of any of those paid sources. The students were prohibited from using phishing emails or phone calls.

At the end of their research, the student groups were to compile their findings into a professional report. Along with the report on the discovered information, the student groups were to describe a possible spear phishing scenario that could be used on the target. At the conclusion of the project the targets were provided with the report and gave the student groups feedback on the organization and completeness of their data, the accuracy of their findings, and the likelihood of success for their spear phishing proposal. This feedback was presented to the students and then incorporated into their grade for the assignment.

The goals of the project were threefold. First, to make the students aware of the extent of the information that was available. Second, to inform the students of the wide array of tools available for conducting OSINT gathering and instruct them on how to use them. And finally, to give the students experience in preparing a written report of a security assessment, deliver it to a customer, and receive feedback on the report.

2.1 Tools

Prior to the assignment, the students were given a lecture detailing many of the current OSINT techniques. For the lecture, we heavily relied upon Michael Bazzell’s “Open Source Intelligence Techniques” [1]. Many of the techniques that we discussed were items that the students had a basic familiarity with, but it was instructive for them to see it all put together in the context of OSINT research, as well as to explore some of the more advanced, and perhaps unconventional, features of these tools.

We began by reviewing basic Google searches and then expanded into more effective ways to search Google, including Google image search and reverse image search. We discussed advanced search operators, including `site`, `linkto`, `filetype`, and `inurl`, as well as time and location filters. We also delved into the Bing search operators `ip` and `LinkFromDomain`. We also introduced Google hacking techniques, such as those described by Johnny Long [8], which makes effective use of these advanced operators. Of course, once something is posted on the internet, it is always available through the many caches and backups that exist, the most useful of which is Google's cache, which allows you to view recently deleted, yet indexed, files. Archive.org can provide historical snapshots of a website over time.

We discussed Facebook graph search and other ways of abusing Facebook in order to search for information, including information that is not supposed to be searchable. For example, Facebook's forgot password page allows you to use a contact's phone number to search, even if this has been disabled by the user. Other sites, such as LinkedIn, Twitter, Reddit, Pinterest, various gaming sites, and Amazon Wish Lists were discussed.[1]

For finding general information about an individual, the students were encouraged to use Reverse Phone Lookup tools, spydialer, spokeo, www.peekyou.com, www.lullar.com, and Zillow, as well as government sites that store information of public record. To find emails and other similar types of information, they were shown TheHarvester and recon-ng. For finding information about internet presence, they were recommended to look at whois, netcraft.com,

[1] The students took this last piece of information to heart, as not long after this lecture a book from my wish list mysteriously appeared on my desk.

and shodan. They were briefly shown Maltego and how to look up document meta-data using tools like exiftool, FOCA, and metagoofil.

Finally, we introduced the students to app.echosec.net, a website that allows you to search social media posts by a geographic area based on any geocoordinates embedded in the posts or pictures associated with the posts. This was, by far, the tool that students were both least familiar with and most impressed and amazed with.

3. PROJECT RESULTS

The students were able to find a large amount of information on each target, some of which will be discussed in this section[2], along with the student's spear phishing proposals and the targets' responses to the student reports.

3.1 Target 1

Much of the available information found for Target 1 was from a defunct website that had last been edited fifteen years ago. Posted on this website were the target's outdated resume, along with information about hobbies and family interests. There was information about the family's involvement in the local PTA, along with details about the target's children, including birth-dates, sports, and many pictures, including ultrasound pictures. While none of the information available was particularly sensitive, it did provide the students with enough information to create a suitable backstory for a phishing attempt. The other danger with this type of information was that it was all available in one particular location and had clearly been neglected in the intervening decade.

3.1.1 Spear Phishing Proposal

The students proposed to use information gleaned about the local PTA from the target's website for their spear phishing email. They theorized that the PTA

[2] Specific details will be either left out or described vaguely, in order to preserve the privacy of the targets.

members would likely pass documents back and forth regarding matters under discussion and that they could embed malware inside an attachment. Also, because the PTA is a public organization, the student group felt that they would be able to make the email convincing, because they would be able to access information about the PTA to hone their attack. Alternatively, the students proposed posing as one of the children's teachers.

3.1.2 Target Response

Target 1 indicated that the information found was accurate and that the spear phishing scenario was plausible. However, Target 1 had suggestions for the student group to improve the overall format of the report and changing its layout and organization. Target 1 also proposed that the students should create a mock-up of a phishing email.

3.2 Target 2

Target 2 had a relatively common name and decades of experience in the security industry. The students encountered many difficulties finding information that was publicly available for this target. Basic work history was found on LinkedIn and a few social interests were located on Facebook. The target's Facebook account was quite difficult to find due to the very common name returning too many results. However, the student group was able to locate the account capitalizing on "features" of the forgot password function on Facebook. The students were able to use another target's name to locate Target 2's Facebook account identification, profile picture, and information about the target's email address and corresponding username.

The students were able to verify that Target 2 had served in the military and discovered a site that verified active duty military status in response to queries based on name, date of service, and birth-date. This validated the information that they had found for the target's birth-date.

3.2.1 Spear Phishing Proposal

The student group devised several proposed lines of attack for Target 2. Their most promising attack scenarios involved using Target 2's identity to attack the target's company. This attack would require more detailed information to be gathered about the specific company, but the information that they had already discovered would have been very convincing if they needed to impersonate Target 2.

3.2.2 Target Response

Target 2 verified that the information uncovered was correct. However, Target 2 requested a more structured and organized report of the information discovered. Target 2 also suggested that the students generate a list of information that they would have targeted with a corresponding list of tools that they planned to use in the attack.

3.3 Target 3

Target 3 proved to be a very difficult target for the students until they discovered the target's mother's Facebook account. The most unexpected find for the students regarding Target 3 was that this target had a degree in the healing arts. The students also discovered information about Target 3's family, place of residence, phone numbers, educational history, work history, and general likes and dislikes. Much of this information was discovered through the target's mother's Facebook page, in addition to public records, familytreenow.com and LinkedIn.

3.3.1 Spear Phishing Proposal

The student group proposed two different scenarios for the spear phishing attack. In the first, they described posing as the healing arts school that the target attended and offering continuing educational classes. In the second, they suggested sending information about a proposed job offer in the nearby large city.

3.3.2 Target Response

Target 3 confirmed that the information that the students had collected was accurate. This student group had included a section of their report entitled "Wild Speculation" which was judged to be largely inaccurate, but indicated that the students were aware of the difference between confirmed information and speculation.

The spear phishing proposals were judged by Target 3 to have moderate to low probability of success. The first email proposed, offering continuing education, would have had low chances of success due to the small nature of the school. The second proposal would have had a greater chance of success except that Target 3 uses a different email address to deal with the work that the target does with the healing arts. The students had not discovered this email address during the course of the assignment.

3.4 Target 4

Target 4 proved to have a larger and more easily accessible internet presence due to the nature of the target's profession. The students were able to accurately find the target's religious and political affiliations, the target's address, and to uncover a court record involving a motor vehicle violation which gave them access to the target's birth-date. The student group was also able to uncover the exact car that the target drives through Google Street View.

3.4.1 Spear Phishing Proposal

There were several promising spear phishing proposals put forth by this student group capitalizing on the public nature of Target 4's profession, which results in a relatively large number of unsolicited, yet legitimate and benign, emails. The team suggested posing as an individual asking Target 4 to either review a website or program that they developed, asking for advice or comments on it, or sending Target 4 a link and asking for the target's opinion about it. The attached program would have embedded malware or the link would be to a malicious site. They also suggested referencing his hobbies in the email, which would add to the legitimacy of the scenario. In addition, they proposed sending an email asking for an endorsement, including a link to their current work, which could redirect the target to a malicious site.

3.4.2 Target Response

Target 4 judged the information included in the report to be accurate, but there were a few small pieces of information which were incorrect. Some of these discrepancies were items that the students had no way of verifying, while others could have been easily confirmed or denied by information that they already had in their possession. For example, they included one link to a high school student's record, not realizing that the student had graduated from high school before the target's confirmed birth-date. Target 4 judged the spear phishing proposal to be very strong. Target 4 admitted to frequently receiving emails like the ones that the students proposed, and while the target vets the links received, admitted that it only takes one mistake for such an attack to be effective.

4. LESSONS LEARNED

At the conclusion of the course the students were asked the following three questions:

1) What did you learn from the OSINT project?

2) What did you find most surprising or eye opening?

3) What changes might you make to your own online habits in light of what you found on your targets?

The students, almost unanimously, were awed by the amount of data that is available on the internet about a given target. They were also impressed that the information they found was linkable to more information. In fact, many students commented on the link-ability of information and how knowing something, like a favorite username, connected a target across the internet and allowed them to be found on otherwise unrelated web applications. They also learned that instead of a search for an individual taking many weeks and requiring specialized skills taking years to hone, anyone with access to the internet and a basic knowledge of the searching tools, particularly powerful Google operators, can find out a lot of information about someone fairly quickly. As one student remarked, “I am aware of how vulnerable I am.”

Many of the students expressed surprise at the amount of information that individuals make available on the internet. It was not just the amount of data available, but the specificity of the data, from children’s ultrasound pictures and birth-dates, to favorite hobbies and drinks, that grabbed their attention. This data was easy for the students to find. Many also commented on the fact that the large volume of data that is available makes it very difficult to remove your internet presence.

The students were not shocked that social media was the biggest leak of information, but they were quick to observe that the information did not always come from the target. First, Facebook itself makes some private information available through its password recovery feature. But, more importantly, and most unsettling to the students, was that other people can put your security at risk. One student commented, “Family members can be the biggest security leak in one’s life.” The students learned that all too often their security is in the hands of others and completely out of their control. [9]

While some students expressed the belief that their online presence was fairly secure, most of them had either made changes to their online activities or planned to make changes in the near future. Many students were conducting a security and privacy review of their social media accounts, locking them down and making them private, as well as filtering the information that is publicly available. Some students said that the project had made them pause before posting information on social media and that they are less likely to post than they were before. Some students also saw the need to not only keep their own information secure, but to make sure that they were not compromising someone else’s information through their social media posts. And, some students, particularly those who were involved with Target 3, felt that they had an obligation to not only teach others how to be more secure or private with their information, but also to confront someone who has compromised the student’s privacy through postings on social media. Besides the big changes to social media, students were also planning to make small, but important changes regarding their internet presence, as one student remarked "the smallest thing can be key to find a new path to more information." In that vein, students planned to hide their birthdays, use fake names for websites that were unimportant, and most notably being careful not to use the same username for every website.

5. CONCLUSION AND FUTURE IMPROVEMENTS

Overall, it is our opinion that the project was a success in achieving the goals that we set out to achieve. As was seen, the students had eye-opening experiences in learning not only how much information was available, but how the various pieces of information, oftentimes small and seemingly insignificant, can be linked together to form a comprehensive picture about a person. One of the other major lessons learned was about how other users can compromise your privacy, even when you have taken measures to lock down your own information. This is a "human" version of what was found in Dodge, et al. in the context of using phishing to teach security awareness, when they observed that "While the application of host and network based security applications can provide some mitigation against malicious activity, there is no static technical defensive measure that can mitigate the threat introduced by human behavior." [4]

We also achieved the goal of teaching the students about what tools are available and how to use them. This can be seen in the comprehensive and accurate profiles that were generated by the student teams. For the final goal, they did get the experience of interacting with the targets, however this goal left the most room for improvement. The quality of interaction can be greatly improved, as we will describe below, along with other areas for improvement that we will implement as we continue to assign this project in coming semesters.

5.1 Rigorous Rubric and More Structured Project Assignment

As this was the first attempt at this project in a classroom environment, it was not clear how the project would progress or what type of information the students would be able to discover. The assignment was given with little structure and no end goals, except to uncover as much as possible about the targets using open sources. Thus, the reports handed in were largely informal and haphazard. In the future, the assignment will be much more structured. The students will be required to present their findings to the target in a professional video-conference using a specific template. The students will also be given more specific suggestions about tools and techniques that they should use and the type of information that they should pursue. In addition, a detailed rubric will be made available to the students ahead of time in order to provide more tangible end goals for the project.

5.2 Spear Phishing Scenario

In the inaugural year of the project the students developed plausible phishing scenarios which were reviewed by the target at the end of the assignment. However, there were several shortcomings with this approach. First, since they were required to describe a spear phishing proposal, the pitfall that many students fell into was to leave the description vague and lacking in detail. The second is that, since the proposals were vague and lacking in detail, the targets were more likely to judge them as ineffective. To address both of these issues, we will have the students create mock emails, including details such as what headers would be in the emails and what links or malware would be included, along with the linked webpages. They will stop short of actually creating malware and sending the email to the target[3], but this should result in a more finished and detailed spear phishing proposal, and may result in more accurate feedback from the targets.

5.3 More Formal Agreement with the Target

The initial targets were all selected based on personal connections with the authors, which will not scale indefinitely. In addition, all the targets were individuals with a security-mindset and involved in the security industry to one extent or another. In the future, this will be expanded to less security-minded individuals with less personal ties to the authors. Our hypothesis is that this would give the

[3] Ideally the students would be able to get to send the email with a link to a benign site, but we are not ready to open that can of worms just yet.

students more varied and interesting results, and would be a beneficial learning experience for the targets as well, making this a classroom exercise as well as an outreach effort. To achieve this, a more formal agreement would be written up explaining the assignment and stipulating the conditions of becoming a target for the students. Since the students are limited to open-sources, permission is not needed from a legal standpoint, but we believe that obtaining formal permission would not only demonstrate high ethical standards, but it will also be a valuable learning experience both for the target and the students.

REFERENCES

[1] M. Bazzell. Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information. CreateSpace Independent Publishing Platform, 2015.

[2] J. Cappos and R. Weiss. Teaching the Security Mindset with Reference Monitors. In Proceedings of ACM Technical Symposium on Computer Science Education (SIGCSE), 2014.

[3] G. Conti and J. Caroland. Embracing the Kobayashi Maru: Why You Should Teach Your Students to Cheat. IEEE Security & Privacy, pages 48–51, 2011.

[4] R. C. Dodge, C. Carver, and A. J. Ferguson. Phishing for User Security Awareness. Computers and Security, 26(1):73–80, 2007.

[5] E. Gava, N. Memon, and D. Britton. Winning Cybersecurity One Challenge at a Time. IEEE Security & Privacy, pages 75–79, 2012.

[6] M. E. Locasto and S. Sinclair. An Experience Report on Undergraduate Cyber-Security Edu- cation and Outreach. In 2nd Annual Conference on Education in Information Security (ACEIS 2009), 2009.

[7] P. Y. Logan and A. Clarkson. Teaching Students to Hack: Curriculum Issues in Information Security. In Proceedings of ACM Technical Symposium on Computer Science Education (SIGCSE), 2005.

[8] J. Long. Google Hacking for Penetration Testers. Syngress, 2004.

[9] B. Schneier. When it Comes to Security, We're Back to Feudalism. http://www.wired.com/2012/11/feudal-security/, 2012. Retrieved 4/13/2016.

[10] R. Weiss, J. Mache, and E. Nilsen. Top 10 Hands-on Cybersecurity Exercises. Computing Sciences in Colleges, 29(1):140–147, 2013.