5-2018

The Development of an Open Source Intelligence Gathering Exercise for Teaching Information Security & Privacy

Jacob A. Young
Bradley University, jayoung@fsmail.bradley.edu

Kerstyn N. Campbell
Bradley University, kncampbell@mail.bradley.edu

Angelica N. Fanti
Bradley University, afanti@mail.bradley.edu

Alex Alicea
Bradley University, aalicea@mail.bradley.edu

Matthew V. Weiss
Bradley University, mvweiss@mail.bradley.edu

See next page for additional authors

Authors
Jacob A. Young, Kerstyn N. Campbell, Angelica N. Fanti, Alex Alicea, Matthew V. Weiss, Justin R. Burkhart, and Matthew R. Braasch

The Development of an Open Source Intelligence Gathering Exercise for Teaching Information Security & Privacy

Jacob A. Young
Bradley University
jayoung@fsmail.bradley.edu

Kerstyn N. Campbell
Bradley University
kncampbell@mail.bradley.edu

Angelica N. Fanti
Bradley University
afanti@mail.bradley.edu

Alex Alicea
Bradley University
aalicea@mail.bradley.edu

Matthew V. Weiss
Bradley University
mvweiss@mail.bradley.edu

Justin R. Burkhart
Bradley University
jburkhart@mail.bradley.edu

Matthew R. Braasch
Bradley University
mbraasch@mail.bradley.edu

ABSTRACT

This research-in-progress paper describes the development of a pedagogical exercise on open source intelligence gathering (OSINT). Exercise materials will include instructions, teaching notes, assessment criteria, and a preconfigured virtual machine (VM), which acts as a local web server. The VM will host multiple websites containing vulnerable information pertinent to a fictitious target organization, in effect creating a capture the flag (CTF) scenario. The exercise will not only teach students how to find public information, but also help students realize the importance of protecting such information. While this exercise is primarily geared towards those pursuing a career in information security, the exercise is appropriate for all students as it shows how personal information could be used against them, as well as their organizations.

Keywords

information gathering, social engineering, security education, information assurance, pedagogy, teaching case

INTRODUCTION

This research-in-progress paper describes the development of a pedagogical exercise on open source intelligence gathering (OSINT) (Bazzell, 2018). Upon completion, exercise materials will include instructions, teaching notes, assessment criteria, and a preconfigured virtual machine (VM), which acts as a local web server. The VM will host multiple websites containing vulnerable information pertinent to TrustUs Community Bank (TCB), a fictitious organization. Students will utilize the VM to complete the exercise in the style of capture the flag (CTF) by attempting to find key pieces of information. Due to the exercise's open-ended nature, instructors will have the option of assigning a variable length of time to complete the exercise. The exercise will not only teach students how to find public information, but also help students realize the importance of protecting such information. While this exercise is primarily geared towards those pursuing a career in information security, the exercise is appropriate for all students as it shows how personal information could be used against them, as well as their organizations. The remainder of this paper will explain the motivation for this project, our development plans, and current progress.

BACKGROUND

Open Source Intelligence Gathering

Security professionals must approach their duties from the perspective of an adversary. Doing so allows for the security posture to be improved by identifying risks and vulnerabilities that need to be resolved. The first step in that process is to analyze the information that an adversary could obtain to increase his or her likelihood of successfully exploiting the target. Such

information is typically gleaned from company websites, social media, public records, and unsuspecting employees. Further, due to the reach of search engine crawlers, OSINT techniques like "Google hacking" often reveal sensitive information that a particular target did not realize was publicly available (Long, Gardner, & Brown, 2015). This critical step in security assessments is presently being taught with an out of class group OSINT exercise.

Current Exercise

The current approach involves having small groups of students each pick a target organization upon which they would like to perform passive OSINT reconnaissance. Over the course of one to two weeks, each group of four to five students competes to collect as much data as possible from various sources. They must then explain how they could use what they found to exploit the organization. For example, the information could be used to plan vishing phone calls, phishing emails, site visits, or crack weak passwords. Since students gather information about actual organizations, the exercise explicitly limits assignment activities to publicly available information that can be found on the Internet. The students are forbidden from having any direct contact with the target organization through any other means, such as in person, by phone, or via email. This is further reinforced through the signing of white hat agreements.

DEVELOPMENT

Motivation

While the current version of the exercise has proven effective, it also comes with challenges and limitations. First, grading such an exercise is a time-consuming process due to the large amount of disparate data to review. Second, although all the information gathered is publicly available, the students are researching real people at real organizations, which is not easily controlled. Third, the findings cannot be used in actual attacks since the target organizations have not authorized them to perform a security assessment. Therefore, the exercise described in this paper will allow for simplified grading, increased control, and the ability to extend its usefulness beyond OSINT.

Overview of Future Exercise

Using a VM, students will be given access to multiple websites that will be used to facilitate OSINT. The VM can either be run locally on individual student machines using VirtualBox or VMware Workstation Player, or hosted on a server. The websites will be populated with information regarding TrustUs Community Bank and its employees. Each student will then attempt to identify and gather as much information as possible about each employee. Instructors will have the option to provide groups with supporting documentation, such as a blank organizational chart that can be used to fill in employee information.

Development Plans

Developing a realistic exercise of this kind is quite labor-intensive. Before undertaking this project, we determined that we first needed to create a lightweight VM to host the fictional websites. After considering multiple approaches, the VM was ultimately built upon the Turnkey LAMP stack (https://turnkeylinux.org) web server. Two websites will utilize the Wordpress (https://www.wordpress.org) content management system (CMS). The first will serve as the website for a TrustUs Community Bank. The other websites will replicate the functionality of popular social media platforms, such as Facebook, LinkedIn, and Twitter.

Second, we would need to create believable content to populate each website with information. In addition to the business information shared on the bank's website, some of the employees will also be members of the social media platforms. A website called Fake Name Generator (https://FakeNameGenerator.com) was used to quickly generate thousands of fictitious personas. A full list of the metadata that can be quickly generated for each user can be seen in Table 1. Once the initial dataset was generated, some personas were modified to become employees of the target organization. Accounts for relatives and friends of the employees will also be included on the social media platforms to increase the realism.

Mother's maiden name	Email Address	Employer	Western Union MTCN
Social Security Number	Username	Occupation	MoneyGram MTCN
Geolocation coordinates	Password	Height	Favorite color
Phone number	Website	Weight	Vehicle
Birthday	Browser user agent	Blood type	GUID
Zodiac sign	Credit Card Number	UPS tracking number	QR Code

Table 1. FakeNameGenerator.com Metadata

Lastly, we would need to develop resources to assist instructors and students. We have already developed background information on the target organization, which will be provided to the students at the beginning of the exercise. By providing instructors with detailed instructions tailored to various skillsets, instructors will have the option of providing their students with as much or as little guidance as they feel is appropriate. Given the volume of information that will be embedded within the VM, instructors will also have flexibility in terms of how long students work on the exercise. It could be as short as a single class session, or conducted over the course of multiple weeks. The grading criteria will be based upon a point system modeled after the flags used for the social engineering CTF competitions held at DEF CON since 2009 (Social-Engineer LLC, 2017).

Future Development

At the time of this submission, we are currently populating each of the three websites with content. While there is still a lot of work to be done, we intend to have the first iteration of this project completed in time to provide a demonstration at the MWAIS conference. Beyond that, we would like to provide the ability for instructors to regenerate the database and grading key so that each time the exercise is conducted it will have unique information. We have also considered hosting the bank's website so that it can be crawled by search engines. This would allow for Google Hacking techniques to be employed to find critical files that have not been properly protected. Suggestions from reviewers, conference attendees, and eventual users of the exercise will help us to further develop this project.

CONCLUSION

This research-in-progress paper described the development of a pedagogical exercise on OSINT. The exercise will provide instructors with a flexible lesson to demonstrate the risks associated with public information. Upon completion of this project, instructors will be able to request access to the exercise materials, which will include instructions, teaching notes, assessment criteria, and the exercise VM. In addition to training future ethical hackers, we feel that this exercise will assist both academics and information security professionals in educating the general population on the dangers of revealing information online.

REFERENCES

Bazzell, M. (2018) Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information (6th ed.). CreateSpace Independent Publishing Platform, North Charleston, South Carolina.

Long, J., Gardner, B. & Brown, J. (2015) Google Hacking for Penetration Testers (3rd ed.). Elsevier, Waltham, MA.

Social-Engineer LLC (2017) *The 2017 Social Engineering Capture the Flag Report*.