Dynamic Application Security Testing - Fuzzing: Brute-Force API Vulnerability Scanning
(Technical Paper)

Understanding Phishing as a Social Engineering Problem: Why Societal Educational Efforts Falls Short (STS Paper)

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Justin Gou

November 1, 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Daniel Graham, Department of Computer Science

Rosanne Vrugtman, Department of Computer Science

Bryn E. Seabrook, Department of Engineering and Society

Prospectus

Introduction

Since the global COVID pandemic hit, as society becomes more reliant on technology than ever, cybercrime has seen a 600% increase ("PurpleSec", 2021). Addressing this problem, however, is no simple task. Whether this means reducing potential risk on publicly accessible systems, such as websites, or taking necessary preventive measures to ensure no employees unknowingly cause a breach, efforts to counteract cyberattacks concurrently improve alongside the issues of cybersecurity. However, as these technologies improve, so do the technologies of the attackers, creating a never-ending cycle of technological improvement. The field of cybersecurity covers an incredibly broad range of topics, all of which are vectors for attackers that must be properly secured for the integrity of information. In this prospectus, the final technical deliverable will focus on the technology of a Dynamic Application Security Testing (DAST) system, used to automate the process of searching for web vulnerabilities on publicly exposed domains. Anything that can be publicly accessed will become a target for attackers, as it is part of the attack surface of a system. Ensuring that all publicly accessible endpoints cannot be exploited is incredibly important in maintaining a systems integrity. In automating this process, publicly exposed web endpoints will be regularly tested for vulnerabilities, thus drastically reducing the risk of a breach from an external attacker.

On another note, in a different field of cybersecurity, social engineering has become a popular technique for attackers. Social engineering is the psychological manipulation of people into performing actions or divulging confidential information. The reason for the upsurge in social engineering attacks is due to the idea that security systems will become more difficult to break into, but people are always vulnerable to voluntarily compromising integrity. Specifically,

phishing, which is the use of fraudulent emails in order to induce individuals into revealing information, has become incredibly popular. In fact, 92% of all malware is delivered via email ("PurpleSec", 2021). To counteract these social engineering attacks, efforts have been made into properly educating employees to avoid falling victim to these attacks. However, as phishing technology improves, it becomes increasingly difficult to discern between fraudulent and valid emails, thus, phishing continues to be an appealing attack for cybercriminals. These phishing attacks are especially evident in the United States, as "74% of organizations in the United States experienced a successful phishing attack. This is 30% higher than the global average, and 14% higher than last year" (Rosenthal, 2021). Naturally, a research question presents itself to search for reasons behind the issue and potential ways to effectively address the problem.

Technical

Robinhood Markets, Inc. exposes many API endpoints to the public that are not regularly scanned for potential web vulnerabilities. These APIs, or application programming interfaces, are used to connect the user applications with the internal computer servers. Large amounts of API endpoints create a significant attack surface for malicious attackers to target. Without regular vulnerability scanning of these publicly accessible endpoints, any vulnerability could pose a large risk on the Robinhood infrastructure, as this simply gives attackers a straightforward way into the system. To address this problem, a proposed solution is to use fuzzing, a method of brute-force dynamic analysis, to automatically test all API endpoints. Prior to research, a Dynamic Application Security Testing (DAST) system had been developed for developers to perform unit testing on endpoints, such as verifying intended behavior. However, the system lacked the ability to scan for unintended behavior, which is often the cause for a security

vulnerability. DAST is implemented as a web API and designed so that other software developers can easily add custom tests and endpoints to be automatically and regularly scanned.

Through research, DAST will be enhanced with various web vulnerability scanning capabilities, including scanning for HTTP smuggling, server-side request forgery, authentication bypass, etc., all of which are very common web application vulnerabilities. Developers can easily add large amounts of endpoints to be scanned through OpenAPI endpoint definition files. Through simply specifying the file, the system will automatically process the API endpoints and perform standard scanning on every single endpoint. Scans will be performed using an open-source, template-based vulnerability scanner, known as Nuclei. This technology allows the researcher to specify HTTP requests to be made and match information based on what was returned by the server, whether that be HTTP status code, HTTP headers, or text in the body of the response. With deployment, DAST will be integrated with Robinhood's internal communication system to report findings to the corresponding developers. This research will allow the internal DAST system to perform automated testing for common web vulnerabilities across hundreds of API endpoints, significantly reducing risk and potential for an external breach.

From the research, the final deliverable will be presented as a technical report detailing how the success of the research enhanced the security of the system.

STS Topic

When Americans were asked what sort of attacks were associated with cybersecurity, 75% of participants listed hackers, password cracking, and malware (Smith, 2017). However, there is a lesser-known field in cybersecurity that is just as important: social engineering. Social

engineering is a field of cybersecurity focused on manipulating humans into performing actions or revealing information. The most common example of social engineering is known as phishing, which is the act of sending a fraudulent email mimicking someone of authority to trick the user into clicking on malicious links or downloading/running malicious scripts. Recently, phishing has become increasingly popular for attackers. “According to the FBI, phishing was the most common type of cybercrime in 2020—and phishing incidents nearly doubled in frequency, from 114,702 incidents in 2019, to 241,324 incidents in 2020” (Rosenthal, 2021). As security systems improve, it becomes increasingly difficult to break into systems. However, as the famous saying goes, a chain is only as strong as its weakest link; the same applies in cybersecurity, where human error is often the weakest link. If a user falls victim to a phishing attack, it compromises the entire system.

Efforts have been put into properly educating Americans into recognizing phishing attacks, however, phishing attacks continue to be successful, causing significant monetary losses for many. According to a study where participants were given a list of fraudulent and genuine websites and asked to differentiate between them, 90% of participants misidentified at least one website, meaning these participants would have been victim to a phishing attack (Dhamija, 2006). Another study showed that “educating novice users using visual cues can partly improve their abilities to detect phishing; however, many novice users still not paying high attention to visual cues when browsing the internet which make them vulnerable to phishing attacks” (Qabajeh, 2018). Some efforts have also been made into creating anti-phishing software, however, as this technology improves, so does phishing technology, essentially causing an endless cat-and-mouse game between attackers and defenders, a common trend in cybersecurity. Some examples of attackers getting around this is by simply avoiding the filters, as attackers can

learn how certain filters work and construct attacks to bypass the filters (Imam, 2017). With all this being said, there does not seem to be a working solution to properly address phishing attacks.

The problem seems to become increasingly complex as technology and phishing techniques improve, despite the efforts to counter the growth of the issue. In this research paper, the technique of wicked problem framing will be used to analyze the research question and potentially raise solutions. The idea behind wicked problem framing was originally inspired by Rittel and Weber's (1973) work on how planning efforts fail. The general idea is that the analysis of the complex problem will help understand the dynamic problem of phishing attacks and reveal indirect and hidden connections. These problems are typically viewed as "impossible" to solve, as the problem is constantly changing or contradictory. In the case of phishing, the problem is always changing, as attackers search for new and more convincing ways to trick users. To list an example, the idea of context-based phishing attacks is a rather new idea that involves attackers gathering personal information about the victim, such as their personal schedule or activities, to make phishing attacks custom and more convincing (Ragucci, 2006). In previous research, the use of wicked problem framing was used to address privacy in the world of cybersecurity, which addresses a rather similar problem, as society falls victim to a huge cybersecurity issue (Alaqra, 2020). Through wicked problem framing, the goal is to encourage the re-organization and re-interpretation of the problem to suggest a technical fix, though merely leaving the grand problem unresolved.

One criticism of wicked problem framing is that the technique fails to find a solution to the problem that will hold for all situations. Through a simple technical fix, large parts of the problem remain unaddressed, hence leaving the problem open for further exploration.

Methodologies

Research Question: How does the average American understand phishing as a social engineering problem?

In the context of phishing, wicked problem framing will be used by first providing background on the problem; why it occurs, who it affects, who causes the problem, current solutions, etc. In doing so, this information will reveal connections between the understanding of phishing by individuals and the attempts to resolve the problem, potentially identifying a reason for the disconnect. In gaining a better understanding of this relationship, it will help reveal why current solutions are ineffective and potentially suggest how to better the current solutions. Understanding this relationship will ultimately allow me to answer my research question, as it would show how Americans currently understand the problem and why educational efforts have failed to further their understanding, causing people to fall victim to such attacks.

Conclusion

This prospectus covers research into a solution to regularly and automatically scan a large amount of API endpoints for web vulnerabilities, as well as research into the social reason for the high success rate of phishing attacks despite proper efforts to counteract them. In both cases, the goal is to reduce the chance of a cyber breach. With a large number of publicly accessible endpoints, this creates a huge attack surface, which is simply a term to describe the quantity of targets for cyberattackers. In introducing regular scanning for common web vulnerabilities, this drastically reduces potential of a breach, as it uncovers underlying vulnerabilities that would otherwise have been uncovered by a cyberattacker. Through this research, the DAST system at

Robinhood Markets has been successfully upgraded to include a user-friendly fuzzing feature for developers to add endpoints to be automatically searched for outstanding vulnerabilities.

On the other hand, another attack vector is to target individuals who have elevated permissions or access to confidential information. Through social engineering, specifically phishing, attackers attempt to gain information or access to a system through the fault of authorized individuals. As security in systems becomes more difficult to crack, attackers naturally sway towards the weakest part of the system: the individuals. Efforts have been placed to properly educate individuals, in hopes of preventing people from falling victim to these traps, however, the success of these efforts is quite minimal. Through this research, using the technique of wicked problem framing, the goal is to obtain an understanding of how Americans understand the problem of phishing and to understand the current state of solutions and why they do not prove as effective as expected.

References

Alaqra, A. S. (2020). *Tinkering the Wicked Problem of Privacy: Design Challenges and Opportunities for Crypto-based Services* (Doctoral dissertation, Karlstads universitet).

Dhamija, R., Tygar, J. D., & Hearst, M. (2006, April). *Why phishing works*. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 581-590).

Imam, F. (2017, August 29). *Anti-Phishing Services: Pros and Cons*. Infosec Resources. Retrieved October 31, 2021, from https://resources.infosecinstitute.com/topic/anti-phishing-services-pros-cons/

Nyeste, P. G., & Mayhorn, C. B. (2010, September). *Training users to counteract phishing*. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 54, No. 23, pp. 1956-1960). Sage CA: Los Angeles, CA: SAGE Publications.

OWASP. (2021). *OWASP Top Ten Web Application Security Risks*. Retrieved October 31, 2021, from https://owasp.org/www-project-top-ten/

Pienta, D., Thatcher, J. B., & Johnston, A. (2020). *Protecting a whale in a sea of phish*. Journal of Information Technology, 35(3), 214-23

PortSwigger. (2021). *What is HTTP request smuggling?* Retrieved October 31, 2021, from https://portswigger.net/web-security/request-smuggling

PurpleSec. (2021, August 6). *2021 Cyber Security Statistics: The Ultimate List Of Stats, Data & Trends*. Retrieved November 23, 2021, from https://purplesec.us/resources/cyber-security-statistics/

Qabajeh, I., Thabtah, F., & Chiclana, F. (2018). *A recent review of conventional vs. automated cybersecurity anti-phishing techniques*. Computer Science Review, 29, 44-55.

Ragucci, James & Robila, Stefan. (2006). *Societal Aspects of Phishing*. International Symposium

on Technology and Society, Proceedings. 1 - 5. 10.1109/ISTAS.2006.4375893.

Rittel, H.W.J. and Webber, M.M. (1973) *Dilemmas in a General Theory of Planning*. Policy Sciences, 4, 155-169. https://doi.org/10.1007/BF01405730

Rosenthal, M. (2021, October 5). *Must-Know Phishing Statistics*. Tessian. Retrieved October 31, 2021, from https://www.tessian.com/blog/phishing-statistics-2020/

Seager, T., Selinger, E., & Wiek, A. (2012). *Sustainable Engineering Science for Resolving Wicked Problems*. Journal of Agricultural and Environmental Ethics, 25(4), 467–484. https://doi.org/10.1007/s10806-011-9342-2

Smith, A. (2017, March 22). *What the Public Knows About Cybersecurity*. Pew Research Center: Internet, Science & Tech. Retrieved November 23, 2021, from https://www.pewresearch.org/internet/2017/03/22/what-the-public-knows-about-cybersecurity/

Zielinska, O. A., Tembe, R., Hong, K. W., Ge, X., Murphy-Hill, E., & Mayhorn, C. B. (2014, September). One phish, two phish, how to avoid the internet phish: Analysis of training strategies to detect phishing emails. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 58, No. 1, pp. 1466-1470). Sage CA: Los Angeles, CA: SAGE Publications.