Another Phishing Training - Role playing as an attacker

A Thesis

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Master of Science
in
Computer Science

by

Saroj Duwal

Bachelor in Science

University of New Orleans, 2019

May, 2022

Table of Contents

Page

List of Figures

FIGURE Page

List of Tables

TABLE Page

Abstract

This is place holder for the abstact sectionLaboris sit sunt nulla elit nisi adipisicing incididunt esse. Occaecat anim ex fugiat aute fugiat enim cillum. Magna do pariatur voluptate cupidatat consequat magna enim occaecat tempor laborum ea aliqua amet. Ex tempor nulla est amet irure dolor veniam occaecat do deserunt cillum ullamco occaecat.

Keywords: keywords, keywords

1. Introduction

The rapid internet adoption in everyday life and the workplace has presented us with new security challenges. Users are more active on the internet, giving attackers more opportunities to attack unsuspecting victims. There are various technical security measures such as firewall, encryption, threat hunting software, and engaging automation to mitigate these challenges. However, studies have shown that the human layer is the weakest link in the security chain [1] and attackers usually start by targeting the most vulnerable link before performing other detrimental attacks. These attacks with human interaction are generally known as "Social Engineering Attacks." Prevalent social engineering attacks such as phishing, pretexting, baiting, quid pro quo, and tailgating use psychological manipulation to trick users into making security mistakes or giving away sensitive information. This thesis will focus on phishing and different detection techniques through our role-playing gameplay.

1.1 What is phishing?

Phishing is one of the most prevalent social engineering attacks in which attackers target users by contacting them through email, telephone, or text message by attackers posing as a legitimate entity [2, 3]. Unfortunately, these attacks are challenging to detect as attackers use the computing infrastructure to fool the victim into doing something but are doing something else while the computing system is working as intended. Due to this, even users with a high-end security system can be victims. An example of such is the infamous case of John Podesta [4], Hilary Clinton's campaign chairman for the 2016 presidential election. The "googlemail.com" in the domain successfully tricked John Podesta and the Clinton campaign's computer help desk to trust the email (See fig:1.1).

Phishing attacks are constantly evolving with different tricks. For example, although Podesta's email shows it was initially generated from "googlemail.com," making it seem like it might be from Google, that might not be true. Attackers can use different spoofing techniques to hide the sender's

identity. Another common trick attackers use (also present in Podesta's email) is to confuse the user with links hidden behind some text/button or confuse the user with redirecting links (example: TinyURL). As a result, the displayed text/link might not be the final destination. Podesta's team's failure to deal with this phishing email led to leaks of more than 11,000 emails which included private conversations with 2016 presidential nominee Hillary Clinton [5].

> *From:* Google <no-reply@accounts.googlemail.com>
> *Date:* March 19, 2016 at 4:34:30 AM EDT
> *To:* [redacted]ta@gmail.com
> *Subject:* *Someone has your password*
>
> Someone has your password
> Hi John
>
> Someone just used your password to try to sign in to your Google Account
> [redacted]@gmail.com.
>
> Details:
> Saturday, 19 March, 8:34:30 UTC
> IP Address: 134.249.139.239
> Location: Ukraine
>
> Google stopped this sign-in attempt. You should change your password
> immediately.
>
> CHANGE PASSWORD <https://bit.ly/1PibSU0>
>
> Best,
> The Gmail Team
> You received this mandatory email service announcement to update you about
> important changes to your Google product or account.

Figure 1.1: Phishing email sent to John Podesta

Successful phishing attacks are very costly to organizations. In 2020, phishing attacks cost US businesses more than $1.8 billion, up from $1.7 billion in 2019 [6]. These attacks can lead to credential/account compromise, giving the attacker access to sensitive information. Attackers may try to use these data for extortion. For example: In 2014, an attack was successful on an invasion

of celebrity iCloud accounts, leading to the embarrassing leaking of nude photos. The leak was initially considered due to a breach on Apple services, but it was later a phishing attack pretending to be Apple and Google and asking them to change their password [7, 8].

Phishing attacks are continuously rising and have doubled since early 2020. In July 2021 alone, APWG saw 260,642 phishing attacks [3]. Additionally, Proofpoint found that more than 75% of organizations faced phishing attacks in 2021 [9]. These uprising trends in attacks have shown some serious need for mitigations for phishing attacks.

1.2 Current Mitigations

The prevention of phishing attacks can be divided into three steps [10]. The first step to stop a phishing attack is preventing the attack from reaching the end-user. We have seen multiple studies on phishing prevention with the help of the machine learning models [11, 12]. Machine learning approaches such as K-nearest, XGBoost, CNN, RCNN, Random forest, etc., are commonly used to detect patterns and generalize phishing attacks. Some of the models have shown promises with more than 90% accuracy—however, a study conducted by What.Hack has shown that only one of the ten anti-phishing tools tested could correctly identify over 90% of phishing websites. That tool also incorrectly identified 42% of legitimate websites as fraudulent [13]. Moreover, attackers are always looking for the best way to bypass these automated systems and develop new techniques if automated systems start flagging their attacks. The evolving nature of phishing attacks calls for an additional layer of security on top of the prevention layer.

If the attacks reach the user, the next step to secure the user is by warning the user. Most modern web browsers and email clients warn users of any suspicious activities they detect. For example, the browser will actively warn users with pop up for probable phishing sites. In addition, browsers provide passive hints to understand links better. Browsers use different shades of white to inform the user about a "fully qualified domain name (FQDN)" (also called absolute domain name), the complete domain name for a specific host on the internet. Figure 1.2 shows a use case for such a hint. Attackers will intentionally have a confusing link to trick users into clicking the

link. For example, although "help.google.com.bubble.com/changepassword" seems like an email from Google, the actual domain is bubble.com. Users can add any subdomain to domains they own, such as help.google.com.bubble.com, which can potentially be used in phishing attacks. Modern email clients provide similar hints for spam emails and notify the users if they can not verify the sender. Active warnings are more effective than passive signs [10]. Still, attackers can easily bypass these warnings by creating new sites and context-aware websites or emails every time they are flagged.

Figure 1.2: Browsers uses different shades to indicate the primary link.

The final step to avoid phishing emails is user training. A study done by Proofpoint shows that 34% of US respondents believe emails with familiar logos are safe [9]. The study indicates a general lack of awareness about phishing campaigns among the general population. As emails are the most frequent method used for phishing attacks, many phishing trainings focus on training users to detect phishing emails. Our gameplay will focus on emails to train the users against phishing attacks.

One of the most common tools to train users is cyber security videos and reading materials. However, Kumaraguru et al. saw that users seldom seek these materials and tend to ignore emails directing them to these materials [14]. In addition, they noticed that most users do not spend much time reading security-related tutorials. This calls for an interactive training program to keep the user focused and engaged during training.

We have seen new and existing training materials incorporating gaming techniques. Gamification has been gaining rapid popularity over the past decade[15]. It increases engagement by incentivizing learners to pay attention and complete activities. We can observe existing training videos incorporating gaming techniques, such as letting users choose the correct option in the mid-

dle of training videos (a mini quiz game) and giving badges after completion. Newer training videos take gamification further and let learners play through various scenarios, make choices and see the rewards or consequences of their decision. For example, Infosec's Choose Your Own Adventure Security Awareness Game[16] has interactive storytelling to keep the user focused till the end of the video.

Gamification has improved the interactivity with the user, but existing training videos fail to cover the technical details that are commonly found in phishing emails. Our gameplay covers various technical aspects commonly found in phishing emails, such as domains, spoofing, and link hiding techniques attackers use to trick users.

1.3 Litearature Review

1.3.1 Serious games

Gaming approaches in education have been used for over a decade[17]. There is a dedicated genre of games (typically online applications), termed as serious games, for using video games to communicate specific information that helps introduce relevant concepts and apply those concepts to solve problems. The primary purpose of these games is to promote learning rather than entertainment. With the help of different game design techniques (rewards, story progression, feedback systems), users are more engaged and immersed while learning. In addition, the virtual world also provides users with a safe space to experiment without real-life consequences.

Serious games are used in many fields such as education, healthcare, and training. For example, "Garfield's Count Me In" [18] helps children in (special education) primary school practice their arithmetic skills. This math game contains different exercises or 'brick,' which form the foundation for a new layer of exercises. The game design help students master the first layer of exercises before moving to the next layer (basic to advanced).

"Killer Flu" [19] (one of many games by "Persuasive Games") is another example of a serious game that attempts to explain how flu mutates and spreads and how challenging it can be for a deadly strain to affect a large population geographically. The game help spread awareness by

making the player take the role of the flu itself, trying to mutate and then spread it in various conditions. Serious games (such as Killer Flu) can place the user as any character in the game to get the idea across. We use a similar concept in our game by placing the player as the attacker (rather than the victim as many existing training materials do.)

Figure 1.3: Garfield's Count Me In

Figure 1.4: Killer Flu

There have been various studies about using games as a practical phishing training module. Hendrix et al. [20] compared the effectiveness of cyber security training tools with some popular games designed for cyber security training and found some positive signs. Prevailing works can be classified into three main categories (discussed below).

1.3.2 Board and card games

There have been studies based on non-computer-based games. For example, Control-Alt-Hack [21] and "smells phishy" [22] are card games aimed to train users against phishing. Both the games show promises in their approaches and teach the user what to be aware of (such as spelling mistakes, phishing links) through their gameplay. After playing the game, they reported higher efficiency and ability to detect phishing emails.

Although both the games show promise in their approach, non-computer-based games have some inherent limitations. The games have a barrier of entry as it requires pre-setup (with the need

for the cards and boards). Furthermore, once the games are deployed, they are permanent, limiting their ability to train and evolve against new phishing attacks. Finally, board and card games fail to communicate the context of the attack and lack examples of where they might really be used. For example, although the game might have a "hiding links" card, the users lack knowledge on how and when it might be used. The limited skills these games provide may not be best suited as an individual training module.

1.3.3 Phishing Link (URL) training

There have been numerous computer games about phishing. However, many studies focus on one common category: training users to verify phishing links. Anti-Phishing Phil [23] is one of the pioneers in this field. Their gameplay puts the user as a fish. The goal of the fish is to grow larger by eating the good bugs (non-phishing links) and avoiding the bait (phishing links). The game has four different levels. Players have to recognize six out of eight URLs to move to the next level.

There are other similar games to Anti-Phishing Phil. Phish Phinder [24] builds upon Anti-Phishing Phil with more user interaction, such as asking for hints and levels. Another example of a link training game is developed by Baral et al. [25]. Baral et al. game story is based on the game scenario of a ballon shooter where the main character has to shoot balloons with legitimate URLs.

All these games have one thing in common: they teach users how to identify legitimate URLs through their gameplay. As URLs are one of the most critical factors while detecting phishing emails, gameplay dedicated to recognizing phishing URLs will serve as a suitable training module. Anti-Phishing Phil results show that their game had a good impact compared to existing training material while detecting legitimate sites.

However, these games might not be suitable as standalone training against phishing. Although URL training games are a sound training module, these games fail to train users on some common tricks seen in phishing attacks. For example, one of the most significant limitations of these games is the lack of context on where the link might appear. As such, attackers can use different link hiding techniques to trick the user into clicking the link. Moreover, attackers use psychological manipulation to trick users into clicking the link by creating a sense of urgency, fake giveaways,

or making it seem like an email from somebody you know to trick people into clicking the link.

1.3.4 Role playing game

"What.Hack" [13] saw the shortcomings of the link-based game and developed gameplay that train the user on links as well as email context. It puts the user as a bank employee required to process emails to acquire contracts and protect their network from cybercriminals. The game approaches the training by having the user role play as a victim and looking at different techniques found in actual attacks. The game's goal is to block phishing attempts and allow legitimate emails. It simulates the harmful effects of phishing by "firing" the employee in-game if they allow too many phishing emails, take too long, or misclassify a significant number of legitimate emails as phishing emails.

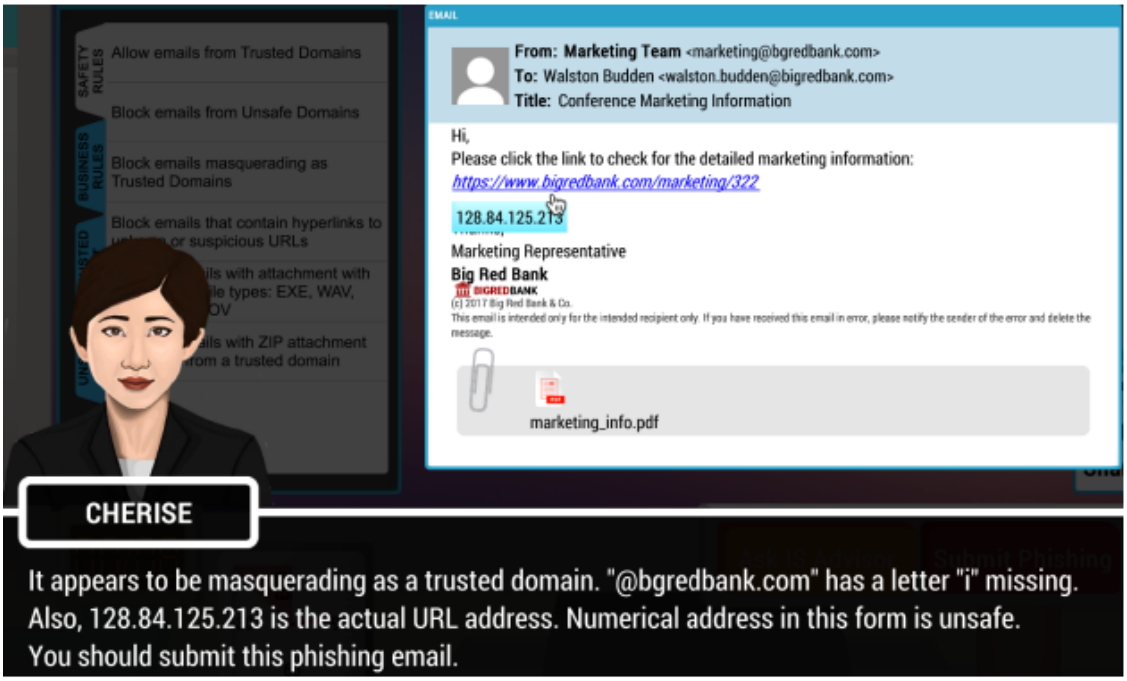

Figure 1.5: What.Hack gameplay

"What.Hack" gameplay incorporates previous studies' different techniques and builds upon them. In addition, its context-based email training module solves one of the drawbacks of link-

based training games. Since players are looking at real emails (both phishing and legitimate), users can better recognize the context of the email and what to look out for in emails to stay protected. Moreover, the emails generated in this game were able to incorporate the primary link-based game goals.

The result from "What.Hack" shows clear improvement regarding link-based games. In comparison to Anti-Phishing Phil, "What.Hack" improved players correctness in identifying phishing emails by 36.7% [13].

1.4 Objective

"What.Hack" clearly demonstrated that role-playing games with contextual emails were more effective than existing gameplays. Unfortunately, we could not find any other significant study that tried to build upon this finding. Therefore, we have developed gameplay inspired by "What.Hack" but approached the role-playing aspect as an attacker instead of a victim.

General phishing training such as videos and reading materials has taught users what to look for as victims. However, we believe looking at the attacker's perspective will help users understand what the attacker might look for while creating a phishing email. This will also help complement the currently available training games.

Our goals for the study can be summarized as:

- Develop a role-playing game to train the users about phishing through an attackers perspective
- Compare our results with the existing study

2. System Description

This is a place holder for system design

References

[1] D. Jampen, G. Gür, T. Sutter, and B. Tellenbach, "Don't click: towards an effective anti-phishing training. a comparative literature review," *Human-centric Computing and Information Sciences*, vol. 10, no. 1, 2020.

[2] KnowBe4, "What is phishing?."

[3] A.-P. W. Group, "Phishing acativity trends report 3rd quarter 2021," tech. rep., Anti-Phishing Working Group, 2021.

[4] "The phishing email that hacked the account of john podesta," Oct 2016.

[5] M. Anderson, "Wikileaks releases more purported emails, bringing total to more than 11,000," Oct 2016.

[6] "Cybercrime statistics: Top threats and costliest scams of 2020."

[7] A. Duke, "5 things to know about the celebrity nude photo hacking scandal," Oct 2014.

[8] "Nude celebrity picture leak looks like phishing or email account hack," Sep 2014.

[9] "2021 state of the phish report," 2021.

[10] I. Vayansky and S. Kumar, "Phishing-challenges and solutions," *Computer Fraud & Security*, vol. 2018, no. 1, p. 15–20, 2018.

[11] R. Yang, K. Zheng, B. Wu, C. Wu, and X. Wang, "Phishing website detection based on deep convolutional neural network and random forest ensemble learning," *Sensors*, vol. 21, no. 24, p. 8281, 2021.

[12] O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, "Machine learning based phishing detection from urls," *Expert Systems with Applications*, vol. 117, pp. 345–357, 2019.

[13] Z. A. Wen, Z. Lin, R. Chen, and E. Andersen, "What.hack," *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019.

[14] P. Kumaraguru, S. Sheng, A. Acquisti, L. F. Cranor, and J. Hong, "Teaching johnny not to fall for phish," *ACM Transactions on Internet Technology*, vol. 10, no. 2, p. 1–31, 2010.

[15] T. Schultz, "Gamification - cybersecurity's turn to play," Dec 2021.

[16] "Choose your own adventure security awareness games," Jan 2022.

[17] L. C. Almeida, "The effect of an educational computer game for the achievement of factual and simple conceptual knowledge acquisition," *Education Research International*, vol. 2012, p. 1–5, 2012.

[18] "Garfields count me in," Sep 2021.

[19] "Killer flu."

[20] M. Hendrix, A. Al-Sherbaz, and V. Bloom, "Game based cyber security training: Are serious games suitable for cyber security training?," *International Journal of Serious Games*, vol. 3, no. 1, 2016.

[21] T. Denning, A. Lerner, A. Shostack, and T. Kohno, "Control-alt-hack," *Proceedings of the 2013 ACM SIGSAC conference on Computer; communications security - CCS '13*, 2013.

[22] M. Baslyman and S. Chiasson, ""smells phishy?": An educational game about online phishing scams," in *2016 APWG Symposium on Electronic Crime Research (eCrime)*, pp. 1–11, 2016.

[23] S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "Anti-phishing phil: the design and evaluation of a game that teaches people not to fall for phish," in *Proceedings of the 3rd symposium on Usable privacy and security*, pp. 88–99, 2007.

[24] G. Misra, N. A. G. Arachchilage, and S. Berkovsky, "Phish phinder: a game design approach to enhance user confidence in mitigating phishing attacks," *arXiv preprint arXiv:1710.06064*, 2017.

[25] G. Baral and N. A. G. Arachchilage, "Building confidence not to be phished through a gamified approach: Conceptualising user's self-efficacy in phishing threat avoidance behaviour," in *2019 cybersecurity and cyberforensics conference (CCC)*, pp. 102–110, IEEE, 2019.