Development and Evaluation of an Anti-Phishing Shooting Game

Heike Dietmann, Tobias Länge, Philipp Matheis, Aleksandra Pawelek, Benjamin Berens, Mattia Mossano, Maxime Veit, Peter Mayer, and Melanie Volkamer

Karlsruhe Institute of Technology
{mattia.mossano,tobias.laenge9,philipp.matheis9,benjamin.berens,maxime.veit,peter.mayer,melanie.volkamer}@kit.edu

Abstract. Background: Games are one type of measure developed to raise security awareness.
Aim: We aim to present the design of a anti-phishing game for public events or for public spaces; as well as insights on how this game is perceived and can be further improved.
Method: We set up a study during a public event in Germany. Participants were observed while interacting with the game and were invited to fill-out a questionnaire before leaving.
Results: Participants left overall positive feedback on the game. However, from both, the observation and the survey, future improvements could be derived.
Conclusion: Our anti-phishing game seems to be a good alternative to classical anti-phishing measures – in particular for public security awareness events. However, further work is required to integrate the findings and then evaluate the game in a controlled study. These are the next steps in our project.

Keywords: Phishing · Gamification · Human Factors

1 Introduction

Internet fraudsters use various strategies to obtain sensitive information from Internet users. A popular method is to send phishing messages, be it emails or through other mediums, asking users, e.g., to make bank transfers, to make (paid) calls, to click on dangerous links, or open dangerous attachments. Phishing is not a new phenomenon (see [20]). Nonetheless, it is still a major threat to this day. The FBI [7] rated phishing as the most widespread cybercrime in 2020 and IBM [11] rated it as the second most costly attack.

Many e-mail providers use technical measures to automatically detect phishing emails. These emails are either not delivered at all or end up directly in the users’ spam or junk folder. Unfortunately, technical measures are not 100%

* This is the extended version of a poster abstract submitted to ACSAC 2022: https://publikationen.bibliothek.kit.edu/1000153327
effective: As attack methods continue to improve, many phishing messages became more difficult to detect and end up in users’ inboxes. One solution could be strengthening the email filter rules. However, despite leading to high accuracy (see [16]), filters are probabilistic and would produce false positives, i.e., remove legitimate emails. Moreover, stronger filters would hardly help in cases of phishing attacks using messaging apps or social networks.

One way to support users is to increase their awareness of the problem and teach them how to distinguish between legitimate and phishing messages. Accordingly, a large number of phishing awareness measures have been developed in recent years, such as texts (e.g., [17,22,24,27]), e-learning platforms (e.g., [13]) and videos (e.g., [9,25]). Besides these more classical awareness measures, games and quizzes were also suggested and evaluated (e.g., in [2,5,10,21,28]). Typically, the idea of these measures is that one consumes them either because it is part of the mandatory security training for employees and/or one is motivated to learn more about security measures. But what about those who are not yet motivated? In many countries, there are several security awareness activities organized by local police stations, local consumer protection agencies and similar organizations at public places - e.g. during the safer internet day which is organized in e.g. Germany and the UK. During such activities usually people are present that can explain the awareness measures and the topics surrounding these measures to interested citizens. One challenge of these awareness measures is to be attractive enough for citizens to stop by, i.e. receiving a leaflet might not be attractive and therefore also not very effective.

The focus of this paper is to present the design of an anti-phishing game and a first study at a public event in Germany. The aim of the study is to gain insights from observing users’ interaction with the game and collecting their feedback on how to improve it.

56 participants tried the game and gave feedback through an anonymous questionnaire. Overall, the game was well received: over 95% of them gave it a positive rating, with such sentences as: “The game is cool!” , “The game is fun!” and “It was fun!”. However, both during the observations as well as from the provided feedback, we derived a list of improvements. Our plan for the continuation of this work comprises the implementation of the collected feedback as well as evaluating the game in a controlled study.

2 Related Work

Phishing definitions: There is no clear definition of phishing. However, generally the focus falls on three main aspects: (1) phishing where attackers deceive users in order to access sensitive information (e.g., passwords, personal data, bank details) using authentic-looking phishing emails or web pages (e.g., in [3,14]); (2) phishing distributing malware when recipients click on links in messages or open attached files (e.g., in [1,23]); (3) phishing where the semantic content of the message induce people to take action, e.g., wiring large sum of money, usually because of persuasion techniques (e.g., in [12,18]) We consider all as phishing.
however, our anti-phishing game addresses only the first two, as the third one is too context dependent (e.g., the CEO of a specific company is impersonated).

Types of Anti-Phishing Interventions: Anti-phishing interventions can be (1) tools or (2) security awareness measures. An overview is provided in Franz et al. [8]. Researchers have created a range of tools (or UI designs for such tools) to provide anti-phishing support to users, e.g., additional security indicators or existing security indicators displayed in different ways (e.g., in [15, 29, 31]). We see these approaches as a supplement to the anti-phishing game that we present.

Various studies evaluated security awareness measures in different formats: videos (e.g., in [9, 25]), games (e.g., in [2, 10, 21]), on-site tutorials (e.g., in [6, 23]) and text-based measures (e.g., in [17, 22, 24, 27, 30]). In our anti-phishing game, we employ an adaptation of the measure in Reinheimer et al. [19] (see Section 3).

Anti-Phishing Games: In Sheng et al. [21], the authors proposed Anti-Phishing Phil, an online anti-phishing game with a focus on URLs, i.e. dangerous links. Users play as Phil, a small fish that need to eat “good” worms to become bigger. The worms “badness” or “goodness” depends on the URL associated with them. Users see the URL associated with a worm and should analyze them carefully before deciding if “eat” them or not. Phil’s father works as a just-in-time source of tips, giving information on how to judge URLs. A demo is available on youtube. While this game may also be a way to attract users during public events, from the design (e.g. sweet small fishes), it is likely to attract kids and families. As attracting people is challenging, it is most effective, if it fits what they like in general. Our target audience are people who like video games with a broader definition of phishing than the one from Sheng et al.

In Canova et al. [4], the authors proposed NoPhish, a serious game divided in ten levels of increasing difficulty. One needs to pass some exercises to continue from one level to the next one. There focus is on phishing emails with dangerous links. While the authors call it serious game, it is more like a e-learning with the levels and exercises. Furthermore, it is not suitable for public events to get in touch with people. Thus, their focus is different from ours and we have a broader phishing definition.

In Wen et al. [28], authors proposed What.Hack, an anti-phishing awareness game in which users are playing as employees of a business having to judge emails. They are pressured through time and action consequences (e.g., refusing too many legitimate emails leads their manager scolding them for being unresponsive). While it is a good alternative to e-leanings in companies, their focus is different from ours as it is not designed to attract people on public events.

3 Proposed Anti-Phishing Game

The anti-phishing game we propose is a first-person shooter, i.e., a type of game centered on combat in first-person perspective. However, as will become

1 https://www.youtube.com/watch?v=c1Es2qzaII
apparent in the descriptions below, a very loose meaning of the term “combat” applies to our proposed game.

In terms of content, the anti-phishing game is based on the security awareness measures proposed by Reinheimer et al. [19], previously evaluated in different formats and shown to significantly increase the ability of readers to distinguish between phishing and legitimate messages. We adapted the content and the material used by them for the evaluation of their measures to fit a shooting game. The game was developed in German.

3.1 Overview

Players are taken to a virtual office. Here, they are positioned in front of a desk with a monitor, a keyboard, and a mouse (see Figure 1). The design idea at the base of the anti-phishing game is that messages fly towards the player, one after the other, and only the phishing ones should be shot at. If the message is legitimate, points are awarded only if the players shoot at the “Legitimate” button on the keyboard (see Figure 1). Additional points are earned the more quickly and accurately a player acts. Accuracy is defined as hitting the malicious part of the message, e.g., hitting the malicious e-mail address, URL, or attachment.

To make the anti-phishing game more entertaining, it has sound and visual effects (e.g., firework explosions for correct decisions). These effects are shown in a small video when the game is not actively used.

The anti-phishing game itself can be played with either a gamepad or with keyboard and mouse. The game is built in the Unity 3D engine. It is available both as a stand-alone application for Windows or as a web application to support various application contexts and operating systems.

The application is based on typical gaming elements. It was developed using an iterative approach while integrating feedback from potential future players.

3.2 Detailed description

Once the game itself is started, there are three options: Rules, Controls, or Play.

Rules. In case someone wants to first know more about phishing, one can select this option. It gives access to three pages containing seven rules to detect phishing messages presented in Reinheimer et al. [19] (see Figure 1).

Controls. When selected, the “Controls” option displays for each input methods (gamepad or keyboard and mouse) how the four actions Aim, Shoot, Look and Pause can be performed.

2 Redacted for for blind review
3 A link to the app will be added only in the final version to preserve anonymity.
Play. This option starts the actual game. One can select to continue with Play training or Start with score mode. “Play training” works as a tutorial and the player sees six pre-selected emails or text messages in consecutive order. The first four show instructions on the correct actions to take. In the last two, the player is free to act. Thus, overall, for those players who do not know much about phishing there are two options to learn more about phishing detection before actually playing the score mode: Reading the 'rules' and 'play training'. Thus, again, depending on how one prefers to consume information, one could select one or the other.

“Start with score mode” starts the main game mode. Ten different emails or text messages randomly selected from a pool of 50 messages and displayed one after the other to the player. These messages are similar to those used by Reinheimer et al. [19] for the evaluation of their phishing awareness measures (to evaluate their awareness measure, they asked their participants to judge several messages – for each they had to decide whether it is a phishing message or not). The messages are a mix of messages potentially sent between friends as well as those one might receive from service providers. For the service provider mails, we decided to use fictional service providers rather than using existing ones to avoid potential legal issues by using the logos of real web services.

In every session played, the anti-phishing game is set to select five phishing messages and five legitimate ones. The players’ task is to decide for each message between phishing and legitimate within a 30 seconds time frame and shoot according to their decision either on the mail or on the “legitimate” button. If the 30 seconds time frame is exceeded before they do, the players get no points.

If several messages are answered correctly in a row, the score received increases through a “combo” system. The amount of points received also depends on the time required to make a decision. Faster decisions are awarded more points.

Leader Board. At the end of the score mode, the players are shown the number of points they have achieved and their position on the leader board. Here, they can either enter a name or continue anonymously.
If used for an awareness event, there are thus at least two different setups: With one screen (game and leader board are displayed sequentially on the same screen) or with two screen (the leader board is shown to further attract people continuously on the second screen).

Overview pages. Afterwards, players see two overview pages with all the messages they judge in their session and the respective decisions. Clicking on a message shows additional information on how one could detect this message as phishing or why this message can be considered as legitimate.

3.3 Application Context

The main purpose of our proposed game is to support organizers of security awareness events in making security awareness more attractive – in particular more attractive for those in favor of video games. Organizers could either stand next to the setup of just be around for questions.

However, the game could also be used by private users (web version) or for use in public spaces, such as museums or show rooms, as stand-alone attractions. In particular the museums and show rooms would introduce new channels to reach out to people in the private context.

Note, after playing through the game in the online version, a new option becomes available in the main menu: A multi-player mode to invite friends and compete directly. To do so one receives a URL to be shared with friends. The idea of this option is to have a new approach to reach out to even more people.

Finally, the game may also be used to check the own performance in detecting phishing messages – in particular for those who believe they already know everything and do not need any anti-phishing awareness.

4 Methodology

This section describes the methodology and the participant recruitment. The goal is to get insights on how people interact with the game and to collect feedback to further improve the game.

4.1 Study Setting

The anti-phishing game was made available at an event in Germany over three days for three hours a day. The event primarily served to provide new residents with information about the city where the authors’ university is located.

Due to the COVID-19 regulations, the event participants were invited on appointment by the city council, so that no large gatherings would happen on site. Note, the authors of this paper were not involved in this process. Thus, they

---

4 Actually, the game was for four months in one museum available for their visitors as part of an exhibition on digitization.
had no information about the people invited. For the same reason, each surface was disinfected after each participant to guarantee the safety of the participants. Moreover, masks were mandatory the whole time. Before playing the game, each participant was required to disinfect their hands before starting interacting with the setup.

The setup of the study consisted of an armchair in front of a large monitor next to a table with a PC, a gamepad, feedback forms, a yellow mailbox and pens (see Figure 2). The station was constantly supervised by one of the authors.

Participants usually arrived individually and reached the station with the anti-phishing game setup on their way out (i.e. after having received welcome information and presents from the city council staff).

4.2 Participant Recruitment, Ethics, and Data Protection

The participants were recruited with convenience sampling. They either approached the researchers themselves or the researchers offered them to try the anti-phishing game. They were told that the anti-phishing game was developed by students (i.e. two of the authors) and that the participants could help to improve it by playing it and providing feedback.

If the station was free, many were immediately interested on their own or quickly agreed to test the anti-phishing game after being approached; few said they didn’t have the time or interest. If the station was occupied, few waited. Participation was entirely voluntary and they could stop at any time. It was also possible to play the game but then not to provide feedback.

All the data was collected on paper and was anonymous (no demographics). We neither had nor requested any information about the people invited to this event.
4.3 Observation

At least one of the authors was present while people played. However, interaction took only placed if questions were asked. Thus, it was entirely up to the people playing whether to read the information about phishing detection (rules) or how to use the controls. They could decide to simply start playing, i.e. getting the messages to decide on.

At the beginning, the researchers noted the time, whether the participants read the rules to detect phishing messages and/or the control instructions. It was also noted if the participants went through training mode or directly to score mode. At the start of score mode, the time was recorded for the second time.

After playing through score mode, the time was recorded for the third time and it was noted whether the participants checked the overview of their answers, i.e. checked in particular the messages which they did not properly decide on. Finally, it was noted with “yes” or “no” whether a feedback sheet was taken, filled out and put in the mailbox. At departure, the time was recorded one final time.

4.4 Feedback Questionnaire

At the end of the interaction with the anti-phishing game, the participants were invited to fill out a questionnaire for evaluation and feedback. On the questionnaire, the anti-phishing game could be rated in terms of design (7-point Likert, from “Very good” to “Very bad”) and understandability (7-point Likert, from “Very understandable” to “Very incomprehensible”). It was also asked to what extent they agreed that the high score is an incentive to play again (7-point Likert, from “I strongly agree” to “I strongly disagree”). The participants were then asked whether they had any suggestions for improvement and, if so, which ones. At the end, they were asked if they would recommend the anti-phishing game to others.

5 Results

5.1 Playthrough Data from Observation

Over three days, a total of 57 individuals participated in the study and were observed. Among the participants, 14 out of 57 (24.6%) read the anti-phishing rules and 7 out of 57 (12.3%) read the information on controls. The majority skipped the first two options and went straight to the play button. Of the 57 participants, 9 (15.8%) chose to practice with a training session first, while 48 (84.2%) started directly in score mode. The average time for an interaction (i.e., starting the anti-phishing game, playing it and moving forward to the questionnaire) was 6.18 minutes, with a minimum time of 2 minutes and a maximum of 15 minutes. The average duration of a playthrough was 3.20 minutes. The first overview page of their results was read by 25 participants (43.9%) while the second one by 21 participants (36.8%).
5.2 Feedback from Participants

A large majority of the people playing the game (52 out of 57) filled out the anonymous questionnaire. When asked to rate the anti-phishing game design, 33 out of 52 participants (63.5%) selected “Good”, 11 (21.2%) “Very good” and 5 (9.6%) “Partially good”. I.e., a total of 49 out of 52 (almost 95%) participants gave the anti-phishing game a positive rating.

When asked how easy it was to understand how to play the anti-phishing game, almost all players (96.2%) felt that the anti-phishing game was generally easy to understand, with such rating as “Very easy to understand”, “Easily understandable” to “Partially understandable”.

Regarding the high score being an incentive to play it again, 36 (69.2%) participants generally agreed, with 13 (25.0%) participants simply agreeing, 3 (5.8%) participants fully agreeing and 20 (38.5%) participants partially agreeing. 45 of the 50 (90.0%) participants said they would recommend the anti-phishing game to others while five (10.0%) said they would not.

Almost half (25 out of 52) of those surveyed also provided feedback. The answers were clustered and the following type of suggestions for improvements were identified:

- Provide more context on the messages to be judged as most senders are unknown. This issue is mainly caused by the design decision to not use real service providers.
- Improve the overview page by focusing on the messages which were judged as legitimate although they are phishes. Currently all messages are presented in the way they appeared during the game together with the information of whether the decision was correct or wrong.
- Improve the overview page by providing general hints how to improve the phishing detection skills (in particular when the wrong answers are above a threshold or for a particular type of phishing).
- Feedback regarding the interaction during shooting: e.g. shooting on everything outside the email could be considered as legitimate; and improving the messages' visibility before it zooms.

6 Discussion and Conclusion

Many people stopped by to try out our game, although they were unaware that there is the possibility to play our anti-phishing game at the event they were invited to. The interaction observation and the participants’ feedback suggest that the proposed anti-phishing game might be a good alternative to well-known awareness measures. It may simply attract people to try it out and learn more about phishing detection from playing the game. The main limitation is that our goal is to find an appropriate measure for security awareness events while in our study setting the main event was not on security at all. Thus, once we have integrated the feedback we received, an evaluation during a security awareness event is planed. This evaluation should also be in a more controlled setting.
Most of the proposed feedback can be easily integrated. The most challenging one is handling the unknown service providers. Reinheimer et al. [19] addressed this issue by explaining to their participants that for the duration of the study, they should assume having accounts for certain service providers. We are currently discussing to change the examples to at least reduce the list of service providers to one or two in order to avoid that people need to read a lot of text regarding service providers before actually starting to play the game. As a consequence, we may reduce the number of messages for both phases, i.e. the 'play training' and the 'score mode'. If this is not improving the situation enough, we are going to reach out to well known service providers and try to get the permission to use their designs for messages including their logos.

References


