Design, Implementation and Evaluation of an Anti-Phishing Education App

Design, Implementatierung und Evaluierung einer Anti-Phishing Education App

Master-Thesis von Clemens Bergmann und Gamze Canova

Januar 2014

TECHNISCHE UNIVERSITÄT DARMSTADT

SECUSO
SECURITY · USABILITY · SOCIETY

Design, Implementation and Evaluation of an Anti-Phishing Education App
Design, Implementatierung und Evaluierung einer Anti-Phishing Education App

Vorgelegte Master-Thesis von Clemens Bergmann und Gamze Canova

1. Gutachten: Professor Dr. Melanie Volkamer
2. Gutachten: Arne Renkema-Padmos

Tag der Einreichung:

Erklärung zur Master-Thesis

Hiermit versichere ich, die vorliegende Master-Thesis ohne Hilfe Dritter nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Darmstadt, den January 30, 2014

(C. Bergmann)

Erklärung zur Master-Thesis

Hiermit versichere ich, die vorliegende Master-Thesis ohne Hilfe Dritter nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Darmstadt, den January 30, 2014

(G. Canova)

Contents

Abstract

Scammers discover the Internet as a convenient place for their criminal activities. For instance, they send Internet users spoofed e-mails or publish fraudulent websites which prompt users to enter their confidential data. This kind of Internet fraud is referred to as phishing. For a victim of a phishing attack the consequences can be of an economic as well as an emotional nature. There exist multiple technical solutions to approach the problem of phishing. Yet, they all cannot guarantee 100% accuracy. Moreover, sometimes security warnings or indicators of such approaches are ignored by end users. For this reason a complementary approach is required. We believe that the increase of security awareness and especially user education about the dangers of the Internet, is a further key strategy to combat phishing. Our master thesis aims at developing a smartphone app, which increases security awareness and educates the user regarding phishing. To increase security awareness the users send themselves a spoofed e-mail right away when starting the app for the first time. This shall exemplary illustrate them how trivial e-mail spoofing is and is intended to increase their motivation and engagement ultimately leading to better knowledge retention. The user education part entails alerts regarding known techniques of attackers and helps them to internalize these by practice and repetition. In detail, our app is realized as a quiz based game which mainly focuses on the detection of phishing URLs. Ultimately, the app should enable the users to achieve the capability of defending themselves against phishing attacks in the future, in case technical solutions should fail. In order to evaluate the effectiveness of the app a user study is conducted. The study outcomes shows that our app in general receives positive feedback from the users and also helps them make better decisions on whether a given URL is legitimate or fraudulent.

Abstract

testtesttest

1 Introduction

Nowadays, a world without Internet is unimaginable for most people in developed countries. As an example, 83% of Germany's population has Internet access [29]. However, it is undeniable that with the benefits of the Internet also come threats. One major issue of today's digitalized world is phishing. Phishing is a term which is referred to for various scenarios and techniques resulting in multiple definitions. s:phishing$_{g}eneralelaboratesonthedifferentphishingtechniquesandscenarios.Inthescopeofthiswork,however,phishingisaf$c

1.1 Consequences of Phishing

Falling for a phishing attack has several consequences for the fooled person as well as for the target company or organization. In the following some of these consequences are briefly illustrated.

Reputational Damage of Target Organization. Moreover, the targeted institutions may sustain a damaged reputation caused by customers falling victim for phishing attacks [53, 77].

Identity and Data Theft of Customer and Target Organization. Phishing is the practice of tricking users to disclose their personal data. That is to say, a possible consequence of falling for a phishing attack is identity theft [42]. With the data unknowingly provided by the victims, the attacker can impersonate them on their behalf. For example, he can do online shopping or transfer money to his account on behalf of his victims. In a corporate scenario the attacker might even gain access to secured systems by attacking an administrator. When the attacker has access to these systems he might be able to collect customer data. Therefore not only users who are subject to phishing attacks can be affected by the attack, but also the institutions, organizations, companies and also their customers.

Financial Loss of Customer and Target Organization. Financial loss can be the result of users' banking accounts being plundered or increased support costs for the targeted institutions due to their customers who fell for an attack [71, 53].

Displeased Customers. Customers who actually became a victim of such a phishing attack will be displeased about the money or account loss and the resulting efforts they have to make in consequence of such an attack. Furthermore, they might tell other people about this unpleasant experience.

Loss of Customer Trust. Ultimately, these victims will lose their trust in the institution targeted by the phisher [77]. Moreover, they might lose confidence in eCommerce operations and the Internet in general.

1.2 Statistics of Phishing

Phishing is also reflected by many statistics of various reports. According to the Anti-Phishing Working Group (APWG) approximately 40,000 unique phishing websites are detected each month [37]. Statistics published by Kaspersky Lab, a well-respected provider for IT security solutions, state that from year 2011-2012 to 2012-2013 the number of attacked users increased by about 87%. While in 2011-2012 the number of users, who were subject to phishing attempts, was 19.9 million, in 2012-2013 the numbers climbed up to 37.3 million. Every day about 100,000 Internet users are victims of phishing attacks, which is twice as many compared to the previous period of 2011-2012. An immense increase can also be observed in the number of unique sources (i.e. IPs) of attacks, which has tripled from 2012 to 2013 [47]. The amount of target institutions also rose. While in 2011 the APWG counted about 500 target institutions, in the fist quarter of 2013 720 target institutions were identified [36]. Finally, RSA and ECM estimate worldwide costs caused by phishing at about $1.5 billion for the year 2012 [71].

Note that according to Moore et al. [58] such statistics might be inherently biased. The problem is, there are several ways to interpret collected data. Hence, every party might assess their data with respect to their interests resulting in diverse statistics. Diversity can also result from setting different foci. Therefore, the reliability of such statistics, including the above mentioned, is questionable. Regardless of the reliability and accuracy of the above mentioned statistics we believe that the education of end users is an important step towards countering phishing. Ultimately, more reliable

and accurate statistics are required in order to evaluate the effectiveness of all the proposed countermeasures against phishing. Next, we discuss the importance of anti-phishing education in general and specifically the need for a mobile app for this purpose.

1.3 Technical Solutions to Counter Phishing

Several technical solutions to counter phishing have already been proposed in literature [69]. As a matter of fact, these techniques are not flawless. This section briefly summarizes these techniques.

Commonly, the phisher sends out a tremendous volume of e-mails to random users which contain links to fake websites. According to Dr. Dobb's, for example, every day 500 million phishing e-mails arrive in user inboxes [72]. There the users are lured to disclose their personal data.

Spam filters. One possible countermeasure to phishing is to filter these e-mails before they even reach the receiver. Various approaches for such spam filters already exist [10, 16, 28], but spam filters also have their drawbacks. First, spam filters might be abused for an invisible form of censorship. Second, a spam filter needs to be installed and applied to the users' e-mail accounts. When using an e-mail service the service provider is generally not allowed to access the users' e-mail without their permission. Therefore, spam filters only apply to users that actively enable them. Third, phishers are constantly improving their techniques to circumvent current spam filters. Consequently, such filters can not assure 100% accuracy. Finally, the strength of the filter controls the amount of false positives and negatives. On the one hand it is possible that phishing e-mails can make it through these filters and might harm the user (false negatives). On the other hand there are legitimate e-mails which may not reach the user (false positives). This might result in a user's loss of confidence, which in turn can result in the user not applying the spam filter anymore [61].

Blacklists. An alternative to protect potentially endangered users from phishing attacks are browsers restricting the access to phishing websites with the aid of so called blacklists. Here, the browsers hold a list of revealed phishing websites. If a requested URL is contained in such a blacklist the access to this website can be restricted or the user can be warned about the phishing website. Several blacklisting approaches have been suggested in literature [51, 90]. The major downside of blacklists is that most of them work reactively. That is to say, there is a certain time frame where phishing websites are active without being blacklisted. In this time frame users can access these website without being warned or restricted and thus are vulnerable to fall for the attack. To resolve this problem multiple dynamic and predictive approaches have been proposed to restrict and/or warn the user from accessing phishing websites [68, 60, 52]. Nevertheless, there is no flawless blacklisting approach, as there are always malicious websites which can bypass such protective systems (false negatives). Also, similar to spam, blacklists may contain false positives, and they can also be abused for censorship. Moreover, these systems require a high effort to maintain, since a regular and realtime update is inevitable in order to make the system effective [69]. Furthermore, there is the weakest link in the security chain: there exist users who ignore security warnings and thus susceptible to phishing and other threats. A field study conducted by Akhawe et al. [**?**] revealed that 10% of Mozilla Firefox' and 25% of Google Chrome's malware and phishing warnings are clicked through, i.e. ignored. Users especially seem to ignore Google Chrome's SSL warning (70.2%). According to the authors this behavior indicates that the user's experience with a security warning has a significant impact on their future behavior. As a matter of fact, in case of disregard of these warnings such systems are unhelpful for those who ignore them.

Visual distinction. A further technical approach against phishing is the automatic visual distinction of phishing websites from legitimate ones. For this purpose it is necessary to identify maliciously duplicated websites mainly based on visual similarities [50]. Various solutions can be found in literature to approach this [17, 18, 89]. However, there is no foolproof solution. In particular, if approaches mainly rely on visual similarities many of them will fail if the phishing website is not a duplicate of the original site. Moreover, phishers will always be able to adapt to sophisticated solutions in order to bypass these security levels. Of course a phisher's adaption to circumvent such solutions must be profitable for him, i.e. it must happen fast enough with reasonable costs. On the other hand, some website providers allow the user to customize some visual elements of the website to distinguish it from faked websites. In the end the techniques are not 100% accurate, i.e. cannot guarantee protection. Additionally, as always the human factor plays a major role here: such techniques will remain of no use for users who keep misunderstanding or ignoring the provided visual indicators.

Takedown. Commonly, hosting providers are urged to take down revealed malicious websites by certain parties, for example: banks, other organizations or specialized takedown companies [57]. The removal of phishing websites is an effective solution, since it implicitly solves the aforementioned problem, where users ignore security warnings: a removed website cannot trick a user into entering sensitive data. Website take downs might raise international and legal issues in case multiple jurisdictions are involved. The phishing website might be located on a different country than the targeted organization. If a takedown company requests the removal from a third country the issue becomes even more complicated. However, according to Moore et. al [**?**] the removal of phishing websites generally follows fairly fast (4-96 hours). They state that system administrators are aware of the phishing problem and take such websites down usually without involving police and court. Although, the takedown follows fairly fast, it is not fast enough. The average life time of a phishing website is 61.69 hours, i.e. 2.5 days [57]. Thus, this approach cannot entirely defeat phishing. During the uptime of the fraudulent website falling for it remains a threat.

In conclusion, there are two major issues of existing techniques.

1. *Accuracy of Technical Solutions.* Technical solutions do not assure 100% accuracy. There is always the potential of false positives and false negatives. Furthermore, attackers can always invent new, more sophisticated deceptions that bypass current prevention systems. The attackers are always first in row, i.e. they create a deception technique and once it is captured and resolved by detection systems, they simply create a new technique or adapt the old one so that it is no longer detected.

2. *User Behavior and Knowledge.* Another major problem with approaches to combat phishing is user behavior. As indicated above users tend to overlook or deliberately ignore security warnings. If the user behavior does not change such approaches will remain unhelpful for those users. The problem is that users primarily use the Internet for purposes such as online shopping, online banking, communicating with relatives and friends etc. Aspects related to security are not of their primary interest when being online. Another factor for overlooking and ignoring these warnings is the lack of security awareness. Some users are just not aware of how easy it is for even unexperienced attackers to duplicate a website and send out fake e-mails. Even if users are aware that there is a certain degree of threat in the Internet, people tend to think the probability that they will face such an attack is very low and that it will not happen to them, until it actually happens to them or to relatives/friends.

In summary, pure technical solutions cannot guarantee 100% protection. In the end the users themselves fall into the attackers' traps. Thus, in our point of view a further key step against phishing is to change the user behavior by increasing the security awareness of the users and offering them a service for education regarding how to defend themselves against such traps. The major question to ask here is whether education will help combat phishing. The opinions on this question are divided. There exist security researchers and experts who argue that user education is pointless [**?**, **?**]. Other sources emphasize the need for increased security awareness and education of the users [**?**, **?**]. It also seems that there exist promising and effective anti-phishing education approaches [**?**, 79], yet with the need for further improvement (cf. section 3). Ultimately, we believe that user education can complement technical solutions and change user behavior. Increased security awareness and knowledge of security issues may first change the user's behavior and attitude towards such technical tools in general. The increased awareness might result in taking security warnings and indications seriously (if they have not already done that before). Additionally, we offer them the opportunity to learn to defend themselves in case such tools fail.

1.4 Anti-Phishing Education on the Smartphone

There are several reasons why we chose to educate users on the smartphone.

Mobility and Size. The main characteristic of a smartphone is that it is mobile and smaller than the well-known desktop computers. As a consequence there is less space on the screen. Many browsers, for example, generally hide their address bars due to the lack of space. With the address bar, the URL and other potential security indicators are hidden. The release of iOS7 featured a key step towards better transparency for the user. The Safari browser displays the host (except

for “www” and “m”) instead of the website title or the URL itself. This might make phishing attacks more difficult to succeed, assumed that users look at and assess this part. Additionally, in portrait mode the host is displayed even when scrolling down the page, i.e. this relevant information is always visible. An interesting question to ask here is whether and how many will follow such an approach. Currently, Android does not support such a functionality. Yet, displaying the host instead of the complete URL or the title is not sufficient to help users detect phishing. There is still a need for URL parsing comprehension for these purposes.

Distraction Caused by Mobility. There is also the fact that users often use their smartphones while on the move, for example, when walking or during a train or a bus ride. These circumstances include distractions from the environment which are unavoidable. These distractions obviously will influence the user’s attentiveness. Hence, smartphone users are even more vulnerable to phishing attacks than the traditional desktop user. This is also indicated by a report of 2011 [11], which says that mobile users are three times more likely to access phishing websites than desktop users. This might also be influenced by the fact that mobile e-mail clients effectively provide no way to check the validity of an incoming e-mail. The potential distraction raises the question whether it has an impact on the user’s education and retentiveness. According to the principles of learning (cf. section 8.6.1) it most likely has an impact on the learning performance. Yet, we believe that our exercise and repetition scheme (cf. section 8.6.1) helps users to internalize the learning content despite slight distractions. For further research it would be interesting to test how significant distractions impact the learning results with our app.

High Number of Smartphone Users. In addition, given that the majority of the people use a smartphone on a regular basis in Spain, Germany, Italy, France and the UK [21], there is a need for the protection of smartphone users.

Benefits of Education on the Smartphone. Educating the user on the smartphone provides two major benefits. First, the user can use the app on the move. Thus, the app is accessible outside of the user’s desktop environment. The app can be used during train or bus rides, while waiting for a friend or while waiting for any other appointment. The app can be started and continued any time as a sideline or just to bridge time. Despite the fact that we mainly aim at motivated users who want to do something about their unknowingness (cf. section 5) we hope that our mobile app might even reach more users. Second, regarding several aspects we believe it is easier to transfer the knowledge of smartphones to desktop computers. For example, the parsing of a URL can be easily transferred from smartphones to desktop computers, as the screen is bigger and the URL is easier to find compared to smartphones. Transferring knowledge from desktop computers to smartphones, on the other hand, raises more complicated issues. The parsing of a URL on a desktop computer, for example, cannot be easily transferred to smartphones. The user needs to know how to access the generally hidden access bar and how to view the complete URL. However, icons or security warnings are probably not easy to transfer in any direction since those differ significantly among devices, versions and browsers.

1.5 Goals

We begin with stating our primary goals of this thesis and describe them in more depth subsequently. The major goals of this thesis is to offer a service which educates users about phishing so that he is less likely to fall for fake webpages. This is an addition and not an alternative to technical solutions to counter phishing. We think that that the following steps are important to achieve this goal.

1. Increasing the users’ security awareness
2. Educate the user with the skills to identify phishing websites.

As already indicated in the previous section the lack of a user’s security awareness seems to be a major issue concerning his security-related behavior. For this reason we want to raise the user’s security awareness by demonstrating within our app that faking e-mail senders and content is very easy. Additionally, we want to make them aware that link texts cannot be trusted. Specifically, the user should be told that links do not necessarily point to the destination URL or website they display. This should happen at the beginning of the app so that the user realizes that the threat of the Internet is prevalent

and that he needs to learn to protect himself. Furthermore, we think it might help if the user practically experiences these aspects and is not only told about it. The practical experience is more likely to increase their engagement and hopefully lead to better knowledge retention and learning performance (cf. Principle of Intensity in section 8.6.1) Moreover, besides technical solutions and increasing user awareness it is also important to give the user information so that they can detect phishing. Therefore, increasing the security awareness is a minor introductory part of the app. Our main focus for the app is the education of the user.

1.6 Outline

This thesis consists of ... main chapters: .... Their purpose is as follows:

Chapter 1 motivates this work. ..

Chapter 2 ...

Chapter 3 ...

...

Chapter ... finally summarizes this work and provides an outlook on future work.

2 Background

The objective of this chapter is to provide the required background knowledge for our further design elaborations. We split this chapter into two parts. The first part deals with the term phishing in general which includes an encompassing definition of phishing, common phishing techniques, attack channels, variations of phishing, the scope of phishing we consider, followed by our definition of phishing we refer to in this work. In the second part we introduce different phishing learning techniques. For better readability and comprehensibility we divided the available learning techniques into the content, i.e. what exactly is the user told, and the used media, i.e. how is the user told something about this topic.

2.1 Phishing in General

This section elaborates on the topic of Phishing in general. Phishing is a term which is referred to for various scenarios and techniques. Consequently, there are different definitions of phishing found in literature. Therefore, we start with a definition that entails all types of phishing. Subsequently, we introduce different phishing techniques, used communication channels and variations of phishing. Finally, we state our scope with respect to the term of phishing and provide our own definition of phishing considered in this work.

2.1.1 Abstract Definition of Phishing

The goal of this work is helping users to distinguish phishing websites from legitimate ones. Since phishing is important in the scope of this work, we are going to define the term first. In fact, phishing is a term that is used by many people in different contexts. Therefore, the following definition is intendedly kept abstract in order to cover all possible scenarios of phishing. At the end of this chapter we will state our concrete definition of phishing considered in this work.

"Phishing is the practice of obtaining confidential information from users and describes a form of identitfy theft. Targeted confidential information includes, but is not limited to, user names, passwords, social security numbers, credit card numbers, account information, and other personal information. " [42]

2.1.2 Phishing Techniques

There are various possibilities how phishers can obtain users' confidential information. In the following we describe phishing techniques that can be distinguished [42, 66].

Deceptive Phishing In deceptive phishing social engineering plays a key role. Here, users are deluded into disclosing their confidential data directly to the phisher without being aware of it. A typical scenario is the unsuspecting user receiving an e-mail from an institution he trusts. In fact, this e-mail is malicious and links to a fake website, where the phisher intends to steal the user's data by capturing the fields the user enters trustfully. Once the phisher obtains the user's data, he is able to impersonate the victim's identity and benefit from this.

Malware-Based Phishing As the term already reveals, malware-based phishing embraces some kind of malicious software running on the user's computer. There are several ways of infecting the user's computer with such malware. Social engineering techniques can be used to convince the user to open malicious e-mail attachments or download malevolent files from a website. Another possibility is to exploit security vulnerabilities. Once the malware resides on the target, various technologies can be utilized to get at the users' data. Keyloggers and screenloggers, for example, track users' data input and send relevant information to a phishing server. Recent research has shown that mobile phone Operating systems are as vulnerable to such attacks as desktop systems. Another way is to make use of so-called web trojans, which appear when users intend to log in. While the user thinks he is logging in on a website of his trust, the entered information is actually transmitted to the phisher.

DNS Hijacking This kind of phishing is also referred to as pharming and includes the manipulation of a system's host file or domain name system. These kinds of tampering result in returning a fraudulent IP address for URL requests and thus

leading the user to a malicious website, even though the URL of a legitimate website had been entered. As a consequence the unaware user enters his credentials into this fake website and the attacker obtains these and can misuse them. These attacks are almost impossible to detect for the user.

Man-in-the-Middle Attack In this form of attack the phisher positions himself between the legitimate website and the user. The user's data input is delivered to the phisher, where he stores the information and then forwards it to the legitimate website. Responses are also forwarded back to the user so that the interference of the phisher does not affect the user's interactions. The gained sensitive information can then be sold or misused in any other way. As everything works as usual for the user, it is very difficult for him to detect such an attack.

Content Injection Content injection phishing refers to the practice of embedding additional harmful content into legitimate websites. This content can, for example, be malvolent code to log users' sensitive information and deliver the input to the phishing server. Well-known types of content injection phishing include, for example, cross-site scripting (XSS). XSS vulnerabilities result from a web application's usage of content from external sources, such as search terms, auctions or user reviews of a product. This type of data supply can be misused and instead of delivering the expected kind of data malicious scripts can be injected.

Search Engine Phishing Other phishing attempts involve search engines. Here, websites with offers for fake low cost products and/or services are legitimately indexed with search engines. Thus, users reach these websites when using the search engine. These offers then lure the user to buy those fake products which in turn leads to the disclosure of their sensitive information, such as the credit card number, to the phisher.

2.1.3 Phishing Attack Channels

Several attack channels exist that can be exploited by phishers to reach their victims. This section introduces possible attack channels [70, 66].

E-Mail E-Mail spoofing is a common way for a phisher to reach his victims. These e-mails usually imitate renowned institutions, organizations, companies or banks the recipients trust. They usually contain a text which will deceive the recepient into doing what it says. Typically a link to a malicious website, whose look and feel is almost identical to the original one, is included. On the malicious website the user is lured to enter his sensitive data which is captured by the phisher. Other alternatives are embedded forms in an email where a user fills in his data directly instead of being forwarded to a website. Finally, sometimes users are even asked to directly send back their confidential data.

SMS An alternative to acquire confidential user data is making use of cell phone text messages. As with e-mails, the text message may contain a link to a fake website, where the user is induced to divulge his sensitive information. The user may also be asked to send back the information directly. Another possibility is to be asked to call back a fraudulent telephone number indicated in the text message. This number usually leads to an automated voice response system which is intended to gain the confidential information from the calling user. This form of phishing is also referred to as smishing, derived from the two terms "SMS" and "phishing".

Instant Messaging In this attack the user receives an instant message from one of his friends. However, the user does not know that his friend's account has been compromised by a phisher. The message usually contains a link to a website asking the user for his instant messenger account information (user name and password). As the link came from a friend many users do not expect something harmful behind this and thus enter their credentials. When the phisher acquires the user's credentials he can continue playing this game with the friends of the newly derived user's instant messaging account, which has just been compromised.

Online Social Networks Using online social networks is similar to using instant messaging services. However, online social networks provide additional valuable information to the phisher. With the aid of user profiles and pinboard entries etc. he can make his baits even more credible and trustworthy. Consequently, the likelihood for his targets to get phished increases.

Voice Phishing A further possibility for a phisher is to send out spoofed e-mails asking the victim to call back the telephone number indicated in the e-mail. To deceive the user, the phisher as usual claims to be from a legitimate and trustworthy institution or organization. The number in the e-mail commonly leads to a voice response system by which the user is tricked to disclose confidential information. Alternatively, the phisher can directly call the user and lure him into divulging his senstitive information. Voice-over-IP (VoIP) further facilitates these kinds of attacks. It makes them easy to execute and inexpensive. Voice Phishing is also referred to as Vishing.

2.1.4 Variations of Phishing

In the course of time different variations of phishing have evolved. This section deals with some of these variations that can be found in literature.

Mass Phishing Here, for example, the phisher sends out a tremendous amount of spoofed e-mails to random users. These e-mails usually link to the phisher's fake website where he tricks his victims to disclose their credentials. In this variation the phisher is not forced or even able to customize the mail to the attacked user. He tries to formulate the mail so that it might fit most users and accepts that some users might not fall for it. The principle of mass attacks is very common and effective, since sending e-mails and setting up websites is almost of no cost and effort nowadays. Even if not all phishing e-mails make it through the spam filters or are not opened: sending out a tremendous amount of spoofed e-mails evidently results in a high amount of victims, not in relative, but in absolute numbers. For example, there exist estimations of 156 million phishing e-mails being sent out daily. Only 16 million of these e-mails win the fight against spam filters. The half of these are opened. 800,000 users of these 8 million e-mail recipients actually click on the contained link and still 80,000 users take the bait according to the estimations [73].

Spear Phishing Unlike mass phishing attacks, spear phishing mainly aims at sensitive information like business secrets, intellectual property or even military secrets. While in mass phishing attacks, spoofed e-mails are sent to millions of random users, spear phishing targets specific individuals resp. groups within organizations to acquire sensitive information. In order to make a deceptive request more credible and personal, knowledge of the targeted individuals and organizations is used. Usually, victims of spear phishing receive e-mails with a malicious attachment and are lured to download it [85]. As sharing documents via e-mail is normal in an organization this does usually not arouse suspicion if the e-mail is from a known person with a legitimate context. This makes spear phishing attacks very hard to detect[85, 40].

Whaling Whaling is a specific form of spear phishing. The target distinguishes whaling from spear phishing. While spear phishing aims at specific individuals or groups within organizations, whaling attacks are after high-level targets, such as senior executives or other leaders in positions of influence [33].

2.1.5 Scope of Phishing in Our Analysis

In the previous sections we have introduced numerous phishing techniques, attack channels as well as phishing variations. As there are more ways of how phishing can be understood, we have to constrain the scope of the term phishing for this work. In literature most of the time phishing is described as the act of gaining sensitive information with the aid of fake websites which trick unsuspecting users into disclosing their credentials [79, 37, 47]. This type of attack is the mostly observed one and is a form of deceptive phishing. For this reason we have decided to focus on deceptive phishing. As aforementioned, phishing websites can be distributed in several ways, including but not limited to e-mail, SMS, or online social networks. Since we set our focus on the analysis of URLs when visiting the website, it does not matter where these links originate from. Any attack channel distributing a link to a fake website will be covered by our approach. However we, and the user should, know that by mere clicking the link to come to the website some information might already be send to the phisher. This includes the validity and activeness of the communication path (e-mail address, phone number, OSN account) and additional information (browser information). We have to accept this because checking the link beforehand is not possible in most situations and also very different, depending on communication path and used software. Finally, there are three variations of phishing we have introduced. Our main focus is the mass phishing attack,

since this is the common one. However, if any spear phishing or whaling attack should involve fake websites, this would be covered by our approach as well. Now that we have restricted our understanding of phishing, the next section provides our concrete definition of phishing for the scope of this thesis.

2.1.6 Our Definition of Phishing

In the following we present a concrete definition which encompasses our understanding of phishing for the scope of this work:

"Phishing is the practice of obtaining confidential information from users and describes a form of identitfy theft. This attack exploits a user's trust rather than system vulnerabilities. More specifically, the user is fooled into believing that he is communicating with a party he trusts and lured into divulging confidential data. This usually happens through phishing websites which look deceptively similar to the originals. Targeted confidential information includes, but is not limited to, user names, passwords, social security numbers, credit card numbers, account information. " [42]

2.2 Phishing Learning Techniques

This section deals with different learning techniques used for phishing education in previous work. For better readability and comprehensibility we divided the related work into two categories: the *content*, i.e. what the user is taught, and the *medium*, i.e. how the user is taught. These two categories can be further divided into several classes. In the following, we are going to provide an overview of these classes, before we provide specific examples of previous work in the next chapter.

2.2.1 Content Classification

The content classification deals with the concrete content of learning which is communicated to the user. The objective of this section is to introduce the different classes of learning content that we identified in previous work.

General Knowledge Transfer Renowned and targeted websites, such as PayPal, eBay or Microsoft provide general and superficial information about phishing [55, 65, 27]. Usually, they deal with questions like what is phishing, how does phishing happen, what the symptoms of phishing are and how to report phishing attempts. Providing the user only with text to the topic of phishing makes it possible to communicate any kind of content, so that the learning objectives can get as complex as one wishes. However, it is likely that users do not like reading too much, especially when it gets complex and difficult to comprehend. Of course the user's willingness to read a lot of complex text about computer security also depends on the user's motivation.

E-Mail Based Knowledge In this class of content, the user is told about the "anatomy" of phishing e-mails [86, 81]. Particularly, they are informed about what kind of hints in an e-mail give indications for a phishing attempt. Indications for a phishing e-mail can be impersonal salutation, requesting personal and confidential information as well as exerting pressure and threatening the user with, for example, account closure. The benefit of detecting phishing attempts before even clicking on a link in an e-mail is that the user would not confirm the existence and active usage of his e-mail address to the phisher. More importantly, the user would not unknowingly download malicious software. The problem with the e-mail based approach is that detecting phishing e-mails by looking at their content becomes more and more difficult [56, 82]. Even if today many phishing e-mails exhibit the obvious characteristic of having no personal salutation or being urgent and threatening we observe a growing number of phishing e-mails that do not make these mistakes and it is likely that these obvious hints will not remain in future.

URL Based Knowledge Sending spoofed e-mails with links to fake websites is a common trick of phishers. On the target website the user is lured to disclosing his credentials. Thus, detecting such fake websites is another possibility to protect oneself against phishing. Here the user is taught to distinguish phishing URLs from legitimate ones [79, 7]. Links to phishing websites are not only distributed by phishing e-mails. Such links can be spread via any communication channel,

such as online social networks or SMS. It is even possible to land on a phishing website by just browsing the web. In these situations knowing how to distinguish phishing URLs from valid ones will help whereas knowledge about phishing e-mails in general will not. The problem with this approach is that as soon as the DNS or host file is attacked even for experts it will get difficult to distinguish a phishing website from the legitimate one (cf. DNS Based Phishing in section 2.1.2). Also, it is unlikely that the user is checking the URL after each click. So the user must develop a strategy when to check the URL (e.g. before entering personal data) and when not.

2.2.2 Medium Classification

The learning medium describes how the learning content is communicated to the user. The objective of this section is to introduce the different classes of learning media that we identified in previous work.

Simple Text The simplest way of transferring knowledge to the user is to just provide text about it. Written language is the most researched kind of medium and generations over generations pedagogues have researched and improved this medium. Alone in this medium there are multiple genres and subgenres which all might be used to transfer knowledge. The main problem is that in modern time many people see the simple text as old fashioned and prefer more interactive learning approaches. Additionally some facts can better be transported with graphics than with text.

Game Based Learning Game based learning tries to communicate the learning content vividly and playfully through a game. Such a game usually has a "background story" and a "mission" the user has to accomplish [79, 86]. The game design is important and depends on the target group. Previous work in the area of phishing, for example, has focused on a fish as starring role in their game, cf. section 3. This might work well for a target group of young age, but will most likely not be appealing to a larger audience. This is also reflected by our phishing survey, cf. section 6.

Quiz Based Learning The quiz based approach is a type of a game which relies on a question-answer cycle without using a specific background story [64]. The advantage of a quiz based approach is that it seems more appropriate for adults and thus will likely be appealing to a larger audience.

Comparison Based Learning A further way to teach users is to let them compare legitimate websites, URLs, or e-mails, with fake ones. Here the user has to decide which of the shown examples are the secure ones [84]. We believe that this form of learning would increase the user awareness, as with this approach one could visualize to the user how difficult it can be to distinguish an original from a fake, especially when they appear almost identical. On the contrary, this way of learning does not reflect the reality, which is a major drawback in our point of view. In real life the user does not have the luxury of chosing between two options, he has only one and has to decide whether this option is trustful or not.

Emdedded Learning The aim of embedded learning is to educate the user on the topic of phishing during his every day life. For this reason the user is sent simulated phishing e-mails. In case the user falls for this simulated phishing attempt he is notified and gets more information regarding phishing and how to protect himself [43, 46]. This approach benefits from the so called "teachable moment". The moment the user realizes that he has almost become a victim to a phishing attack, he will be highly motivated to prevent this happening again and thus be highly receptive for input related to this topic. The teachable moment itself will not suffice to make the user stay on and consult the educational page, though. For example, there was a study in Germany to assess the effectiveness of CMU/APWG's landing page. During the study the authours found that people just closed the window immediately after or shortly after landing on the educational page without reading on, because they thought they were on the wrong website and were not aware that they landed on an educational site. Even if with an effective landing page, the missing positive feedback is a major flaw of this strategy in our opinion. The user is only notified in case of a mistake and not in case he has successfully discarded the simulated phishing e-mail. A further problem is raised with the implementation of such an approach. Legal issues will arise when sending simulated phishing e-mails which claim to come from a reputable vendor, such as an online shop.

We believe that a mixture of the game and quiz based approach is the best way to go in order to create an incentive for the users. Regarding the content which will be communicated to the user we decided to focus on detecting phishing URLs for the reasons explored in this section.

3 Related Work

In the previous section we introduced the different classes of learning contents and communication media. Furthermore, we have decided to go for the game/quiz based approach while focusing on URL based knowledge. This section summarizes concrete examples of previous work on anti-phishing education. For each class we previously introduced at least one example is illustrated. We especially elaborate on the game and URL based approach as this is the path we take. Moreover, we state in which way our work is to be expands on previous work.

3.1 Game and URL Based Approaches

Several work has been done in this area [6, 7]. However, most approaches are similar to Anti-Phishing Phil [79] the approach we focus on in this section. Anti-Phishing Phil is a game based approach focusing on URL based knowledge. We will extensively discuss this game since the approach we envision resembles Anti-Phishing Phil the most.

Game Design and Rules The three objectives of this game are the following: (1) learn to detect phishing URLs, (2) learn where to look for indications in browsers for trustworthy/untrustworthy websites, and (3) learn to use search engines to find legitimate websites. The major focus, however, is set on the detection of phishing URLs. The main character of the game is a little fish, named Phil, who has to grow to a big fish by eating worms. These worms can either be good, i.e. real worms, or bad, i.e. fake worms, with which fishers try to hook the fish off the sea. Good worms of the game are associated with URLs of legitimate websites, while bad worms are associated with the URLs of phishing websites. Phil's task is to feed on legitimate URLs only. He must reject phishing URLs to grow to a big and healthy fish. The game consists of four rounds in total, each round takes two minutes. For correct actions Phil is rewarded with a certain amount of points. If Phil falsely rejects a legitimate URL, he is slightly penalized by having the time left decremented for a couple of seconds. However, if Phil eats a phishing URL he is severely penalized by losing one of three lives. In this way, the authors try to simulate the extent of the real world effects of their behavior, i.e. in reality rejecting a valid URL is not as bad as trusting a phishing URL. Each round the focus of deception techniques is shifted and phishing URLs get more difficult to identify. In the first round the users get introduced to IP address URLs. The second round mainly deals with deceptive subdomain URLs, where the brand name occurs in the subdomain. In the third round, the users are taught about similar and deceptive domains. In the last round finally, the user has to deal with all kinds of deceptions he has dealt with so far. The information material provided to the user is delivered by so called training messages. Anti-Phishing Phil features four kinds of training messages. First, the user gets direct feedback during the game, whether the answer he has given is correct or not and why. Second, the user has the possibility to receive help in case he needs it. In this case, Phil's experienced father will give a tip. Third, at the end of each round a score sheet is displayed, which summarizes the user's answers, whether they were correct or wrong, and why they were correct or wrong. Finally, there are anti-phishing tips in-between the rounds.

Game Evaluation To evaluate the effectiveness of the game the authors conducted a between-subjects experiment with three training conditions, represented by three groups: (1) existing training material, e.g. from eBay or Microsoft, (2) anti-phishing tutorials which were created based on the game, and (3) the game itself. Each group had to decide on ten websites (in total 20) about their authenticity before and after the training step. The results showed that the participants in the game condition performed better than those in the other two conditions.

Positive Sides All in all, we believe that the approach is a good step towards user education and features many good aspects. In the first place, the game based approach is an attractive incentive for the user to be educated. Furthermore, the training messages are kept short and simple. Finally, the training messages, especially the ones of help during the game and the score sheet after each round are very valuable. Due to time restrictions we cannot integrate those kind of messages.

Downsides and Our Contributions On the other hand, we believe that this approach has some flaws and thus is not optimal for user education. Even though using a fish as main character for this game is a funny idea, we do not think that this is an appropriate solution for adults. This is also reflected by the results of our phishing survey, cf. section 6. Therefore,

we will not use a fish as our main character. Our approach will rather be a combination of a game, which includes lives and points, and a quiz, where users are required to answer questions directly, without any background story. As aforementioned, the training messages are simple and easy to understand, however, we are afraid that the phrasing is too vague. For example, for IP address URLs Anti-Phishing Phil displays the following alert message to the user: "Don't trust URLs with all numbers in the front". For subdomain attacks the following wording is used "Don't be fooled by the word ebay.com in there, this site belongs to ttps.us". These kind of messages are susceptible to misinterpretation. Another downside we see, which is ultimately related to the vague formulations, is that the user is not concretely explained how he has to parse the URL in order to make healthy decisions on the authenticity of such. Here again he is only told that the most important part of the URL is between the "https://" and "/" and that the name of the website is the text right before the "/". In our point of view this is a imprecise phrasing and there is a lack of emphasizing the importance of the domain, which we do in our solution. Finally, the game does not cover some spoofing techniques, which are still relevant in our opinion and thus are covered by us (cf. section 7.4.1). For example, the difference between http and https is not introduced, as well as the fact that https websites can also be phishing websites. Furthermore, the game does not explicitly mention that the domain name, the host or even the entire URL can be part of the path to fool the user. Finally, there are different ways of making use of deceptive domains, which were not explicitly covered by Anti-Phishing Phil. For example, homograph attacks (cf. section 7.4.1), typos and scrambled letters should be distinguished in order to exemplify how mean and hard to detect such URL spoofing techniques can be.

3.2 Game/Quiz and E-Mail Based Approaches

Anti-Phishing Phyllis [86] is a game based approach and focuses on teaching the user to detect a variety of phishing traps in e-mails. These include, for example, fake links, attractive offers, urgent requests, or malicious attachments. The main character of this game is a fish named Phyllis. Phyllis has to decide whether potential traps (marked with red bubbles) in an e-mail he is shown are real phishing traps or are harmless by disarming or ignoring them. The playing user gets hints during the game and direct feedback on his actions. Another quiz and e-mail based approach is provided by SonicWALL [81]. The user is shown e-mails consecutively and has to decide whether the displayed e-mail is legitimate or not. The user does not receive direct feedback on his decisions. At the end he receives an overview of the answers he has given and whether they were correct. If the user wants to know why his answer was correct or wrong he has to click on a link to get this information. As aforementioned, teaching users to detect phishing e-mails before even giving them the possibility to land on phishing websites has the advantage that they do not confirm the activeness of their e-mail address, and more importantly, do not have the chance to accidentally download malicious software. However, as phishing e-mails become more and more sophisticated, i.e. convincing and credible, and since phishing websites are also reachable via other communication channels, such as SMS, online social networks or just surfing in the Internet, we decided against the e-mail based approach.

3.3 General Knowledge Transfer With Embedded Learning

There are several proposals in literature for embedded learning [44, 46, 1]. One of these is a solution proposed by Jansson et al. where simulated phishing e-mails with links to fake websites or malicious download attachments are sent out to users [44]. The moment a user falls for a trap of these simulated e-mails he receives a notification informing him that he could have fallen for a real phishing attempt. Also, the e-mail includes a link to a website with a training program with general information and tips on how to detect phishing and malicious attachments. After consulting the training program the user is asked to complete a questionnaire in order to verify whether he understood the content of the training program. A very similar approach, the so called PhishGuru [46], is proposed by Kumaraguru. Another possibility is to leave out the step where simulated phishing e-mails are sent to users. Instead actual phishing e-mails are utilized. For example, the APWG and Carnegie University's CyLab Usable Privacy and Security Laboratory (CUPS) work on the project "Phishing Education Landing Page" [35]. The moment a user clicks on a link of a real phishing website which has already been taken down, i.e. the moment the user behaves riskily, he is redirected to the anti-phishing landing page. There he is told that he had almost become a victim of phishing and provided with educational material to this topic. Finally, there is

an approach where the intervention does not happen after clicking on a dangerous link, but while surfing [1]. When the user lands on a blacklisted phishing website and is about to disclose his sensitive data (i.e. presses the submit button) the system interferes: the user is warned and given tips on how to detect phishing websites (e.g. abstract information on the detection of spoofed URLs). All of these solutions benefit from the so called teachable moment, the moment the users place themselves at risk by either clicking on a link in a (simulated) phishing e-mail or by submitting sensitive information to blacklisted phishing websites. This moment of risk presents a teachable moment for those who almost fell for such a trap. For this reason giving the warnings, hints, and training to the user in this moment will most likely result in higher motivation and retention so that the tips are more likely to help avoid similar dangers in future. A downside of these approaches is that they do only give negative feedback to the user. Consequently, the user is not "rewarded" when he rejects to click on a phishing link or to submit data on a phishing website, which is an important thing to do in our view. Moreover, the amount of information provided on such an educational website should be reasonable, i.e. the user should not be flooded with information. Otherwise he will not retain or even consult everything. To overcome these issues, a reasonable consideration for future work might be to combine embedded learning with another approach, for example, playing an educational game. In this way positive feedback can be included and the information can be transferred bit by bit to the user. For example, the website the user is redirected to might contain just the most important information, just enough to motivate the user to click on the provided link to an educational game, for instance our app, in case the user is interested in gaining in-depth insight on this topic. For now, we do not follow this approach since the step of sending simulated phishing e-mails to users raises legal issues.

3.4 Comparison and URL Based Approach

Symantec offers a "race to stay safe" [84], where the user is shown two snapshots of two websites, while one website is a fake and the other is a legitimate one. Within very short time the user has to compare the snapshots and decide which way is the safe one to go. The focus of this training is set on the URL and address bar. We believe that such an approach is likely to increase the user awareness of how deceptively similar phishing websites can be to the original ones. However, the approach of comparing two websites is not realistic enough, since the user does not have two websites and does not have the option to choose between them in reality. This is why we did not decide for the comparison based approach. However, adding time pressure to our approach, i.e. simulating a real life situation, is an aspect which might be worth to consider for future work.

3.5 General Knowledge Transfer With Quizzes

There exist online quizzes where the user is asked general questions to the topic of phishing [9, 64]. The design of these online quizzes is based on the association of phishing with fishing. That is to say, here again a fish is the main character of the quiz, which we do not find appropriate for adult users. Moreover, the number and variety of the questions asked in these quizzes are very restricted. Even if the examples of the quiz based approaches are not optimal for user education, we think that this approach is the most appropriate one for adults as target group.

3.6 Further Game Based Approaches On Other Computer Security Topics

Besides the proposals for user education on the specific topic of phishing, there exist a variety of other approaches aiming at educating the everyday user on general or other specific topics of computer security. Auction Hero [19], for example, is a simulation game which covers different topics of computer security, amongst other phishing. Its aim is to help users make more secure decisions in the Internet by modeling their regular Internet behavior. Real life is simulated by making security a secondary goal of the game, like it usually is the case with end users. The primary goal of the user, who is a trader, is to build and sell robots, and earn enough money and reputation to ultimately become an "Auction Hero". As in reality, the trader has to pay attention to various security risks, such as weak account passwords, out-dated antivirus software as well as phishing. Phishing, in particular, is dealt with as follows: the playing user receives e-mails within the game. While some of them are legitimate others are not (for example, an e-mail saying that the user has won an auction for an item on which he has never bid). The e-mails include links to websites where the user is asked to enter his in-game login data. An ultimate consequence of disclosing data to a phishing website is that the user will suffer loss of money and

reputation. Also, an explanatory warning will be displayed. The user is taught about typical characteristics of phishing, potential consequences of falling for them, and how to deal with phishing attempts. This approach has the major benefit of simulating actual online behavior and thus provides a realistic context for the user. Additionally, the user does not only learn about phishing, but other security related aspects, such as having strong passwords and keeping antivirus software up-to-date. On the other, our aim is not to develop a multi purpose application, but rather to focus on phishing attacks in detail. Targeting multiple security related topics means that the taught content needs to be constrained, otherwise retaining the learnt content will be difficult. Additionally, as security is a secondary goal and multiple security aspects are reflected in the game the user will have to play it for a long term in order to obtain helpful knowledge, especially for the purposes of detecting phishing attacks. Therefore, we want to focus on phishing attacks in detail instead of giving the user an overview of security topics which have to be considered when using the Internet. There exists a variety of other online games and quizzes covering miscellaneous topics of computer security. "Mission Laptop Security", for example, is a quiz based approach where the user's mission is to transport a laptop to a specific destination in a secure manner [63]. During his trip, the user is asked several questions about how to act in different situations. The mission can only be completed if the user gives enough correct answers. Another game covers the topic of network security [62]. Hackers, represented by little red men, are surrounding the user's WLAN area. By clicking on a hacker man a question appears. When the user gives the correct answer, this specific hacker man disappears. When the user gives an incorrect answer all red men come closer to the user's computer in the center of the WLAN area. Others have diverged from computer based games and rather suggested a physical card game primarily intended to increase the users' awareness of needs and challenges related to computer security in general [25]. A further interesting approach is to change sides. That is to say, the user does not play the role of an unknowing user who has to defend himself from attacker's. Instead he is the attacker and his goal is to profit from criminal activities. Data Dealer [24], for instance, provides a game dealing with the topic of data privacy, data abuse and surveillance. The player's, i.e. attacker's, goal is to collect private data of friends, neighbours or any other person, sell the collected data over the black market and set up companies.

4 Scope of Our Approach

This chapter elaborates on the determination of our scope to educate people about phishing. Here, we gather all of our choices and the corresponding reasoning for our defintion of our scope. These include the various classes of phishing and phishing learning techniques that we mentioned in section 2 as well as new aspects. Subsequently, we summarize the system requirements and what assumptions we had to make. Finally, the limitations of our work are stated.

4.1 Coverage

Deceptive Phishing as Phishing Technique Within the scope of this work we focus on deceptive phishing. In particular, we target the detection of phishing websites resp. phishing URLs.

Several Attack Channels As mentioned above, we concentrate on the detection of phishing URLs. Phishing websites can be reached in several ways. Links to fake websites are usually distributed via e-mails, instant messages, or online social networks. Moreover, they can be spread via SMS or even phone calls. Ultimately, a phishing website can also be reached by just surfing in the Internet. By teaching on identifying spoofed URLs, our approach covers all attack channels as long as the user is tricked into divulging sensitive information on a phishing website.

Mass Phishing as Variation of Phishing We cover mass phishing, as already stated in section 2.1.4. However, the URL checking can be applied in case of any variant, as long as the attack includes a website which lures the user to type in his credentials.

Game and Quiz Based Learning as Communication Medium As discussed in section 2.2.2, we have decided to develop a quiz game to create an incentive for the users and at the same time reach a large audience.

URL Based Knowledge as Learning Content As argued in section 2.2.1 we decided to educate the users about phishing based on URLs. We believe that URL based knowledge gives the most reliable hint regarding its "origin", i.e. whether a URL in fact belongs to a legitimate website or not. Additional we had a look at the phishing URLs provided by PhishTank [67]. The majority of these URLs were not or only loosely related to the attacked website. If the users would be aware of the importance of the URL and were able to interpret it the phishers would put more effort in forging valid-looking URLs. Obviously, there are enough users falling for these primitive attacks. Therefore, we think that it is important to inform the users about the significance of URLs and to teach them how to interpret those.

After Click URL Analysis The analysis of a URL can follow before or after clicking on a link (if a link is involved), i.e. with a URL preview option or directly in the address bar. Analyzing the URL before clicking on it brings several benefits:

1. *No Malicious Download* If a spoofed URL is detected before clicking on a link the potential download of malicious software can be avoided.
2. *Phisher Obtains No Information* Recognizing the spoofed URL before visiting the website prevents the phisher from obtaining any information of the user. Such information include, for example, the activeness and validity of the user's e-mail address.

On the other hand, the before click scenario has severe downsides:

1. *Redirects Not Recognizable* Many links contain redirects. Such redirects, which can be malicious, are not recognizable before the click.
2. *Unavailability of URL Preview Functionality* The stock e-mail client of Android does not provide the functionality of previewing the destination URL. Here, the only way to preview the URL to apply a long press to the link, copy it into the clipboard, paste it somewhere else and then view it. Then, after the analysis the URL has to be sent to the browser. As this is involves too much effort, it is likely that no user will follow such a suggestion.
3. *Deception With URL Preview* There are other e-mail clients which provide options to display the destination URL of a link. Yet, we believe that this should not be communicated to the user for two reasons. First, we are elaborating

on a general approach that does not rely on third party e-mail clients besides the default one, which is pre-installed and comes with the device itself. Second, and most importantly, this functionality has the potential to mislead the user. A severe downside of the URL preview is that the end of the preview is cut in case the URL is too long. Well-crafted URLs might thus look legitimate even though they are not because the most important part of the URL, i.e. the actual domain, was cut out. For example, the subdomains of the URL can be long and well-crafted so that a legitimate looking subdomain is exactly at the end of the preview. This will cause the user to think that the subdomain at the end of the preview is the domain of the URL. Ultimately, the user will trust this URL even he should not.

4. *Users Like Clicking* Clicking on links is practical and convenient. As a matter of fact, users like clicking on links. Hence, we cannot hinder users from clicking on links.

The after click scenario does not exhibit all these drawbacks which is why we chose to follow this approach. On the other hand, this scenario might suffer from potential malicious downloads and providing information to the phisher (benefits of before click scenario). If a user confirms his e-mail address to the phisher by clicking on a link, further attacks towards this e-mail address are likely to follow. Also, the pure request and displaying of a phishing website might provide additional information to the phisher or even expose the user to attacks. For now, we consider this as future work, as there is no possibility in our target scenario to detect the real target of a link before clicking on it.

Considered Browser As a matter of fact, the user is only taught general browser skills, which can be transferred to any other browser. Nevertheless, when the user gets browser screenshots, for example, we made use of the Android standard browser to be sure that most users are familiar with the pictures they are shown.

4.2 System Requirements

In the following we list the system requirements which need to be met to install and play with the final app.

Android Currently the mobile phone market is split between two major competitors. Android (81.0%) and iOS (12.9%) [41]. We have decided to develop an app for the Android operation system as there are more users and we believe we have greater freedom here compared to an iOS app. Additionally, Android is an open operating system and imposes less requirements and barriers that allow us a quick development and publication of our app [13, 12].

Version Our initial intention was to develop an Android app for version 4.0 and upward. However, during the app development we have encountered that about 24% of all Android users still have Android 2.3.3 to 2.3.7 [5]. For this reason we have decided to modify the code so that these users can also install and use our app.

Samsung Galaxy S3 or S4 Actually, it is not necessarily required to install the app on Samsung Galaxy S3 or S4 devices. Yet, we did not have the possibility to test our app (design and functionality) on different devices. The above mentioned devices are those we tested and used for our final user study.

4.3 Assumptions

We have to make some assumptions about the user's system. If one or more of these are not met the user, disregarding of his skills, might not able to detect when he is a target of an attack.

Secure DNS We have to assume that DNS is not under the control of the attacker. Our approach is to show the user how to identify phishing attempts by analyzing the URL of the shown page. In fact, this is of no use if the phisher can control the DNS system of the user. Therefore, we assume local host files and all used DNS servers as untouched by the attacker.

Secure Smartphone We imply that the user's system is in a secure state. This means that the attacker is not able to, for example, exploit browser vulnerabilities, replace the browser or read the user's input.

Secure SSL For a man-in-the-middle it is possible to intercept sent or received data and to collect the user data directly without notice. Therefore, we warn the user against entering personal data on non-HTTPS pages. But even in HTTPS environments we can not be sure of the server's identity if SSL is not secure. We are aware of events showing that there are a multitude of SSL providers which, intentionally or not, fail to award certificates to legitimate users only. When the certificates cannot be trusted, SSL cannot be considered as secure and, ultimately, the user has no practical possiblity to detect a man-in-the-middle attack.

Malware Sending e-mails with malicious attachments or links to malicious downloads are a form of deceptive phishing. In our approach we assume that the attacker does not lure the user into downloading such malicious software, which captures the user's confidential data. We focus on attacks where the user is actively encouraged to provide his sensitive data himself. Yet, we believe that preventing users from being lured into downloading malware is an important aspect. Therefore, in future work one should consider how our approach can be expanded with regard to this problem.

4.4 Limitations of Our Approach

In addition to the general assumptions pointed out in the previous section, there are limitations of our approach resulting from the chosen target group for our app. As described in section 5, users from our target audience are no computer experts and have neither time nor are they willing to analyze the shown website thoroughly before entering data. Therefore, we do not intend to tell the users about possible attacks that only experienced user might find.

Cross-Site Scripting Cross-Site Scripting (XSS) is an attack where the attacker enters code, such as a form, into a legitimate webpage. For a later viewer of this page this form seems to be legitimate content of the attacked webpage and he might be lured to enter personal data in this area. Depending on the attacked webpage this cannot be detected by the user. This is a vulnerability of the attacked webpage and can be prevented by checking user input. Therefore, we think that preventing this attack is in the responsibility of the website owner.

URL Hiding Techniques Most modern mobile phones have small screens. Therefore, most browser hide away the URL bar to increase the viewport when the user browses webpages. The URL bar is shown only when the user scrolls up beyond the top of the webpage. There is a possible attack where the attacker prevents the user from scrolling all the way up and instead displays a fake URL bar with a fake URL. A problem that the attacker faces is that mobile browsers look very diffently and this must be reflected by the fake URL bar - a tedious and difficult task. Moreover, attackers do not need such sophisticated attacks because enough users fall for the simple attacks these days. We think this is the reason why this kind of attack has not yet been observed in the wild and the reason why we do not consider it as well.

5 Target Group

This chapter deals with the target group we want to reach with our app. After defining our target audience we briefly explain how it can be projected to the German population.

5.1 Target Group definition

In this section we want to describe the targeted users for our app. The main condition that must be met is that they can learn something from our app. That means they are skilled enough to use the app and not too skilled so that they already know everything that the app tells them. In detail this is modeled by the following conditions.

Attackability The first precondition is that all our users must meet is that they are possible targets for phishing. This means they must use the Internet. They also should use the Internet often enough and have a common trust in the web so that they are in general willing to enter their personal data [26].

Android Users The second precondition is that the users should use an android smartphone. Our evaluation shows that the app is also usable by iOS users but they are not the target group because they cannot use the app on a regular basis.

Language The informative parts of our app are texts and they are written in German. This means the target user should be able to read German texts.

Motivation The distribution plan for this app is to put it on the Google Play Store and hope that users download and install it. Therefore, the target user must be willing to learn something about the Internet. According to a study from the DIVSI [26], some Internet users are so sure about their knowledge that they are not willing to learn anything else. We will not be able to reach these users.

5.2 Projection to Population

After we have decided what our target group is, we wanted to make sure that we do not exclude too many of the potential users with these preconditions. In fact, the app can only be useful and successful if a reasonable audience is covered. To prove this we looked at an extensive survey done by SINUS-Instituts Heidelberg on behalf of Deutschen Instituts für Vertrauen und Sicherheit im Internet (DIVSI) [26]. In this survey the authors first looked at 60 persons in detail and found seven types of Internet users that are depicted in Figure 1. Thereafter, they tried to apply the findings to the whole German population by interviewing 2,000 representative persons. Figure 2 illustrates the percentage share of the seven Internet types in Germany. We tried to match our preconditions to these groups. In the following we present our evaluation of each of the groups.

Digital Souveräne This group moves naturally on the Internet and is therefore exposed to phishing. They also often use smartphones. We rule them out because they think that they already know the problems of the Internet and hence they would reject a training offer anyways. In fact, this group will likely never download the app.

Effizienzorientierte Performer This group matches our preconditions because they are using the Internet as well as smartphones. In contrast to the previous group, they are interested in learning something new and see their own learning as an investment in the future. To target this group we should show that you can learn something from this app.

Unbekümmerte Hedonisten This group is also native in the digital worlds but in contrast to the before mentioned groups are not aware of the problems and frauds therein. When they are aware of the problems they seek to secure themselves with automated software instead of concerning themselves with it. Therefore, they are not motivated to use our app.

Postmaterielle Skeptiker This group is interested in the Internet and uses it frequently. On the other hand they are aware that there are problems and frauds. As they are interested in information on the Internet especially from official sources they might download our app. To target this group we should clearly state that this app is from an university.

Figure 1: Internet Milleus as defined by DIVSI [26]

Verantwortungsbedachte Etablierte This group is online regulary and also uses smartphones. They are especially interested in using protection software and actively search information on the Internet. The users of this group do not think that they could protect themselves from the dangers of the Internet and actively seek to change this. Therefore they most likely will appreciate the app. To target this group we should clearly state that this app helps the user to protect himself.

Ordnungsfordernde Internet-Laien These users are using the Internet rarely. Because of this they are particularly careful when using the Internet and normally do not enter personal data. Therefore, it is not likely that they will use the app. Besides, they usually do not have smartphones.

Internetferne Verunsicherte These users do not use the Internet. Therefore, they are not exposed to phishing threats.

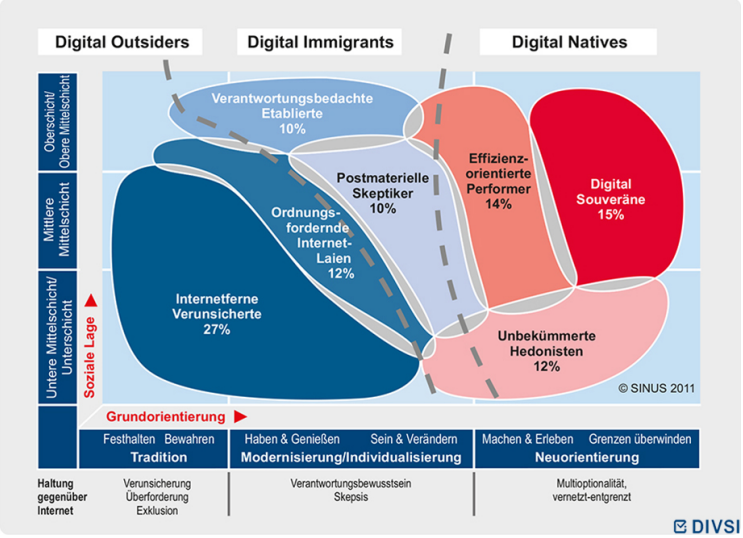


Figure 2: Internet Milleus as defined by DIVSI

In conclusion, we consider *Verantwortungsbedachte Etablierte* (10%), *Postmaterielle Skeptiker* (10%), *Effizienzorientierte Performer* (14%). In total these are 34% of the german population.

6 Phishing Survey

Before elaborating on the app design we ran a phishing survey of which we illustrate the main objectives in this chapter. To the best of our knowledge there do not exist other surveys which resemble ours and additionally were conducted in Germany. Furthermore, it provides details regarding the contents of the survey and finally presents the results and evaluates the questionnaire and how they have influenced our further app elaborations.

6.1 Main Objectives

Our main objectives of this survey were twofold:

1. *Awareness and Knowledge* One goal of the survey was to comprehend what exactly Internet users understand under phishing. With a Likert scale we furthermore tried to figure out how they evaluate their knowledge on Internet security.

2. *Preferences of Users* Another purpose of the survey was to get an idea of the users' preferences with regard to an educational app. For example, they were asked whether they found a quiz based game appropriate for learning purposes.

6.2 Survey Details

This section provides details about our questionnaire, how we distributed it and how we filtered the surveys in order to consider our target group for the results and evaluation. The complete survey form can be consulted in Appendix C.

6.2.1 Questionnaire

In the following we present the structure of our questionnaire and the function of each section.

1. *General Information* In this section the participant is asked to provide information regarding his gender, age, his professional qualification as well as his field of study or work. The main purpose of this section is to exclude participants which do not fit into our target group.

2. *Internet Usage* Here, the participant is asked how often he uses the Internet, whether he owns a smartphone and which applications he uses on his desktop computer and which ones he uses on his smartphone. This section is intended to give us an overview of the users' Internet usage and helps us to exclude participants who do not fit into our target group.

3. *Self-Assessment* In this part of the survey, the participant has to indicate how much he agrees to the presented statements with the aid of a Likert scale. The statements mainly refer to their self-assessment regarding their knowledge about Internet security. For example, they have to assess, whether they think they have enough knowledge to avoid the dangers of the Internet or whether they think it is easy for them to distinguish legitimate e-mails from fake ones. This section is partially based on Likert scale statements used by DIVSI[26].

4. *Phishing* Here, the participant gets concrete questions to the topic of phishing. In particular, he is asked which services and which user information are endangered by phishing attacks. This section purposes to find out what the participants know and think about phishing.

5. *Anti-Phishing App* This section asks the user for his preferences regarding an anti-phishing education app. With the aid of a Likert scale he is requested to assess, for example, whether the would like having a game with a fish, whether he finds a text-based approach appropriate, or whether he would have fun with a question-answer (quiz) game.

6. *Further Survey Progress* In this part of the survey the user can provide his e-mail address in case he wants to get information about the further progress of the survey or if he would like to test the app.

6.2.2 Distribution

In total 251 persons participated in our survey. We set up an online survey as well as asked students to fill out our printed survey. In the following we briefly explain our distribution process.

Printed Survey To reach participants for our printed survey we contacted multiple professors and asked them whether we could have 10 minutes of their lecture time to have their students fill out our printed survey. Moreover, we asked our friends and relatives whether they can ask their friends, colleagues or customers to fill out the questionnaire.

Online Survey The online survey was mainly distributed digitally. We contacted our friends and asked them to participate in the survey. We also demanded to forward the link to their friends to reach a wide range of people.

6.2.3 Select Targeted Participants for Evaluation

Table 1 summarizes what kind of answers we used in order to exclude participants from the survey who do not fit into our target group.

In the succeeding section we present and evaluate the results of the study. With the filtering above we had 169 remaining participants matching our target group and thus were considered for the evaluation.

6.3 Results and Evaluation

The study yielded interesting results which should be considered when designing an anti-phishing education app, either for this as well as for future work. This section outlines the results of the study.

General Information The gender ratio of our survey participants was more or less balanced. 40.83% of the users were female and 56.80% of them were male. The remaining 2.37% did not indicate any gender. The average age of our participants is 27.59, the youngest participants are 19 years old, the oldest are 63 years old. Most of the survey participators, 48.52%, obtained a university degree. 24.85% of them do not have any professional qualification (yet). 17.75% did an apprenticeship and the remaining participants had a master craftsman certificate or did not indicate any professional qualification in the survey.

High Rate of Android Users The majority of the participants were Android users. In total about 60% of the study participants use an Android smartphone. The remaining 40% are iOS users. This result roughly relates to the general market share for these platforms and additionally supports our decision for the implementation of an Android application.

High Internet Usage Frequency 51.48% of the users are online several times a day. Another 30.18% indicated that they are online even constantly. As a consequence, over 80% of the survey participants are frequently online. This is depicted in Figure 3. Being online is always connected with being attackable and vulnerable to dangers of the Internet, such as phishing attacks, while the extent of the vulnerability of course also depends on other factors(e.g. the expertise of the person being online). However, the more often a user is on the Internet, the more likely it is that he will experience a phishing attempt.

Usage Distribution of Internet Applications Figure 4 and Figure 5 summarize the usage distribution of Internet applications on a desktop computer and on smartphones. Almost all participants, 99.41%, make use of e-mails on their desktop computer. 88.76% of the smartphone owners use their e-mail application on the smartphone, which is still a high percentage. As we have previously mentioned in section 2.1.3, e-mail is a common attack channel for phishing attempts. Consequently, all users of e-mail applications, including webmail, on mobile phones are potentially endangered by phishing attacks. The same threat of phishing applies to participants using browsers. About 80% of all considered participants make use of desktop or smartphone browsers. Furthermore, it is conspicuous that banking is far less used on smartphones compared to desktop computers. While about 74.56% of the participants make use of online banking on the desktop computer, only 26.63% use it on their smartphones, which is still a quarter of the participants. The question to ask here is if these users utilize the browser for online banking or if they use dedicated apps provided by their bank.

Question	Filtering
Age	We consider all adults ranging from 18 - 65 years.
Gender	We do not exclude any gender.
Professional qualification	The participant does not have to exhibit a specific professional qualification to be considered for the results and evaluation.
Field of study/work	Students, employees or employers in the field of computer science or electrical engineering are filtered out as they do not belong to our target group.
Frequency of Internet usage	Participants who have indicated “rarely” as the answer to this question do not belong to our target group and thus are filtered out.
Used Internet applications	The listed applications include, for example, browser, e-mail, shopping as well as banking. Any service of the Internet is potentially endangered by phishing. For this reason we do not use this question to sort out participants.
Owning a smartphone	With the app we particularly target smartphoner owners. For this reason participants who do not own any kind of smartphone are filtered out.
Used smartphone applications in the Internet	The listed applications include, for example, browser, e-mail, shopping as well as banking. Any service of the Internet, especially on a smartphone, is potentially endangered by phishing. For this reason we do not use this question to exclude participants.
Number of received commercial e-mails per week	We do not filter out any participant with this question.
Number of received e-mails asking for personal data	We do not sort out any participant with this question.
User reads up on topics related to dangers in the Internet	Participants who have chosen “no” as their answer are filtered out. We specifically target users who are interested in getting safer on the Internet. As the participants, who have indicated “no”, do not seem to have any interest in doing so, they will most likely do not show interest in our app. For this reason we regard them as not belonging to our target group and exclude them from the analysis and evaluation.
Section to self-assessment regaring their knowledge about Internet security	We do not filter out participants based on their selection in this section.
Section to questions related to phishing itself	We do not filter out participants based on their selection in this section.
Section to preferences for an anti-phishing education app	We do not filter out participants based on their selection in this section.

Table 1: Filtering rules for the phishing survey

Regardless of the answer to this particular question, these users might be more likely to react to phishing e-mails on their smartphone, which claim to come from their bank, compared to other users who manage their financial arrangements on a desktop computer and thus are less likely to access a phishing website, cf. section 1.4. To sum it up, all the listed categories of applications are used by the participants, on their smartphones as well as on their desktop computers. For this reason, all of these application categories should be reflected in the choice of the example URLs for the final app. For future work, one could argue to put the focus on URLs from specific categories (also those which were not considered for the study), depending on the usage distribution.

Self-Assessment - Knowledge to avoid dangers of Internet 18.34% of the participants think that they have enough knowledge to avoid the dangers of the internet. Further 45.56% agree with the statement and only about 13% disagree or strongly disagree with this statement. As a consequence the majority of the participants were quite confident that they could avoid the security-related risks raised by the Internet.

Self-Assessment - Distinguish legitimate from illegitimate e-mails 87.23% of the participants think that they can easily distinguish legitimate e-mails from fake ones strongly. Only about 8% of the participants did not agree or strongly disagreed with this statement. This arouses the suspicion that the users are not aware of how easy it is to spoof the "from" field of an e-mail, or to create credible message contents which in fact may persuade the receiver to be trustful.

Self-Assessment - Trust in e-mails from known parties The majority of the participants trust e-mails which come from persons they know. Approximately 20% strongly agreed and approximately 57% agreed with this statement. Only about 2% strongly disagreed and approximately 5% of the participants disagreed with this statement. This again shows that most of the participants are not aware that spoofing the "from field" of an e-mail is easy to achieve. These users are likely to react to e-mails which claim to be, for example, from friends. Such e-mails may actually contain links to the download of malware or malicious websites.

Self-Assessment - Internet security is only related to financial applications With a Likert scale the participants had to indicate how much they agreed with the following statement: "Internet security is only related to financial applications". The answers to this statement showed that the majority of the participants are aware that security related issues on the Internet do not solely concern financial applications. 49.7% of the users strongly disagreed with this statement and another 24.26% disagreed. Only about 10% of the participators agreed or strongly agreed with this statement and about 14% indicated "neither nor" as an answer. Even though most users seem to be aware that Internet risks do not only concern financial applications, the ones who are not aware that phishing, for example, can also occur in online social networks, should be enlightened about this. To do this our plan was to display the consequences of falling for a certain phishing website (phishing URL). In this way, the user could have learnt what his loss could have been, projected to a real-life scenario. This would have contributed to his awareness that security issues in the Internet, in this particular case phishing, are not necessarily related to financial loss only. Due to lack of time we could not realize this approach. However, it is something that should be considered in future work.

Services endangered by phishing Figure 6 summarizes the results for this question. All in all, we can observe that the participants agree that phishing can actually occur related to any service. The users agree (97.04%) that especially the e-mail service is endangered by phishing. Also they see the browser (70.41%), online banking (83.43%) as well as social networks (74.56%) as endangered. Still about 40% consider various media (audio and video) services as well as online games as endangered. These services are in fact not targeted as often as other services on the Internet, however they are potential targets and should be communicated to the user with the aid of the choice of the URLs to decide on, for example.

Data endangered by phishing Figure 7 outlines the results of this question and illustrates that the participants agree that any of the listed kinds of data is potentially endangered by phishing attacks. 90.53% of the participants expressed that login data is endangered by phishing. About 89% agree that credit card information as well as personal data is endangered, too. Ultimately, 76.33% of the participants consider PINs and TANs endangered. Consequently, there does not seem to be a major necessity in enlightening users in this area.

Preferences for an education app This section influenced how our app is designed. There were three results that we took into consideration. First, most of the users (50.8%) stated clearly against an app that uses a fish as the main character. Second, most of the users either voted for a quiz based game (51.5%) or did not care (33.1%). The minority of the participants voted against a quiz based game (13%). Finally, 40.8% of the users were neutral regarding text based learning programs. For these reasons, we could confirm our desired approach of a quiz based game, with introductory parts that also contain text, cf. section 8. The other results from this section remain to be considered for future work. These include aspects, such as dividing the program into exercise and test mode.

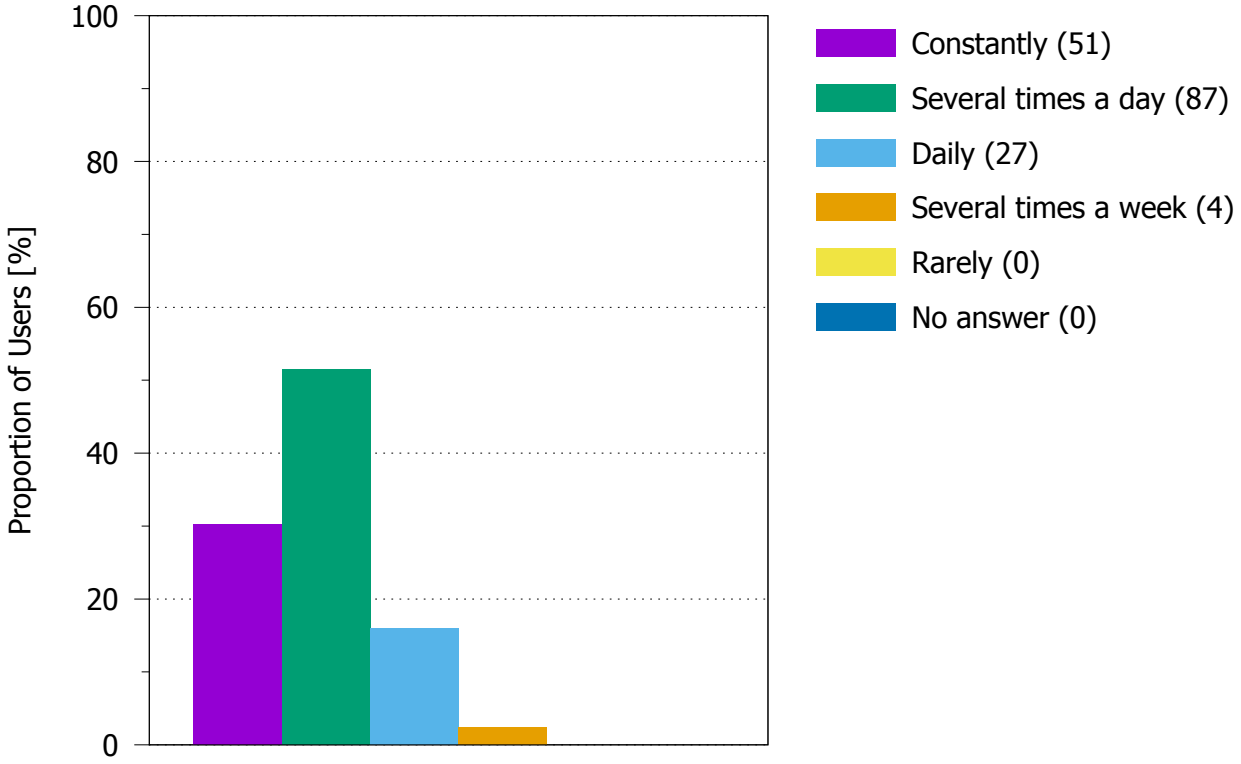


Figure 3: Frequency of Internet Usage

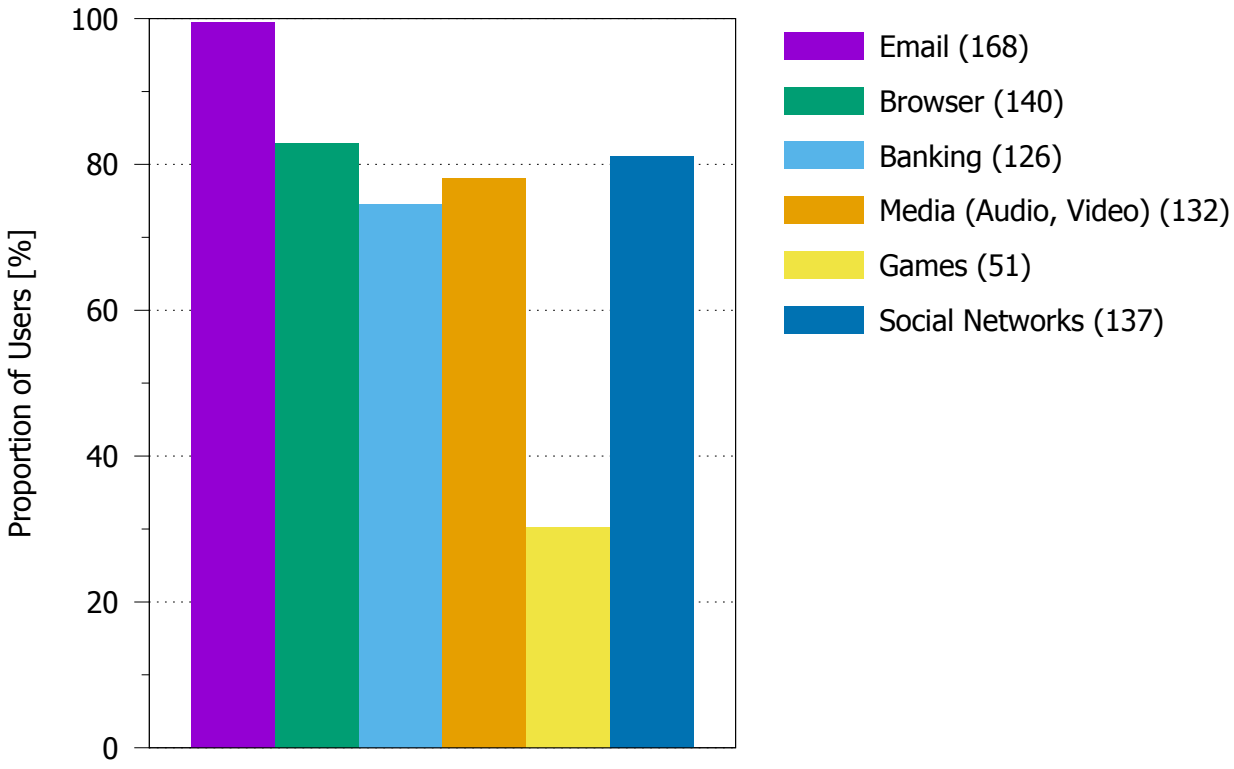


Figure 4: Usage of Internet Applications on Desktop Computers

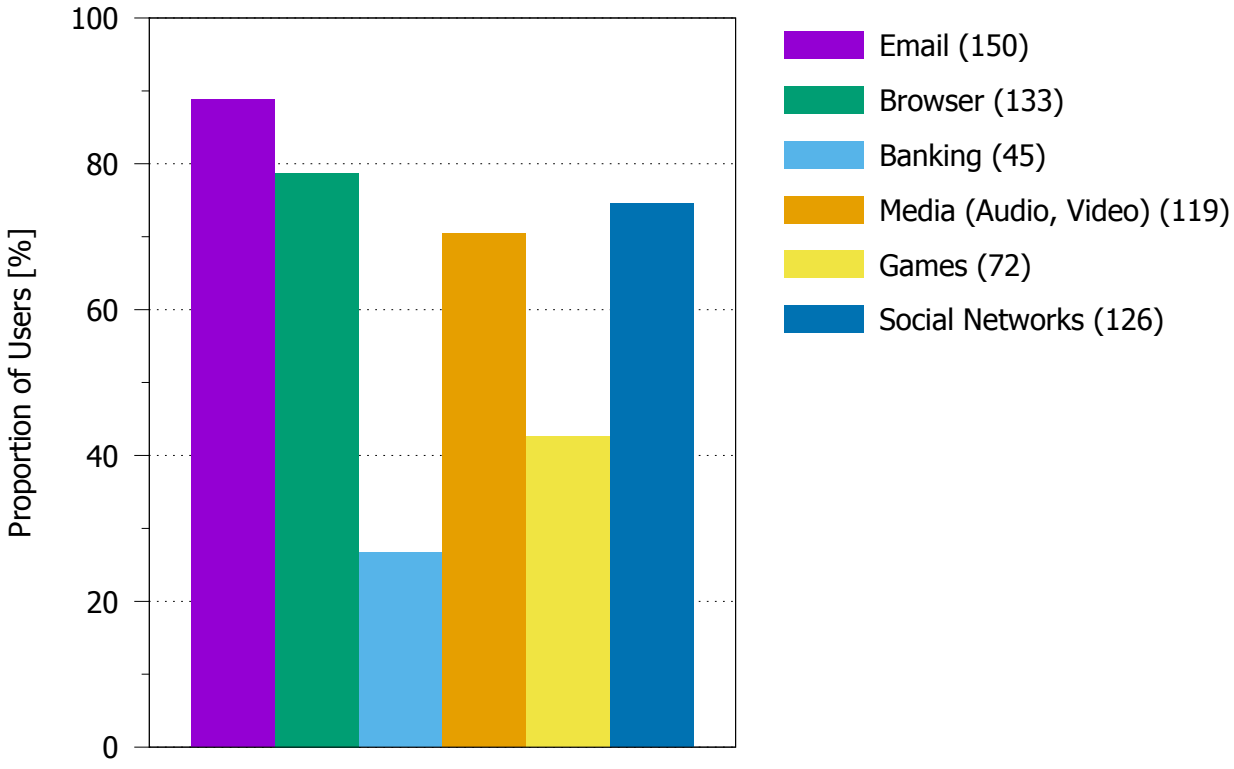


Figure 5: Usage of Internet Applications on Smartphones

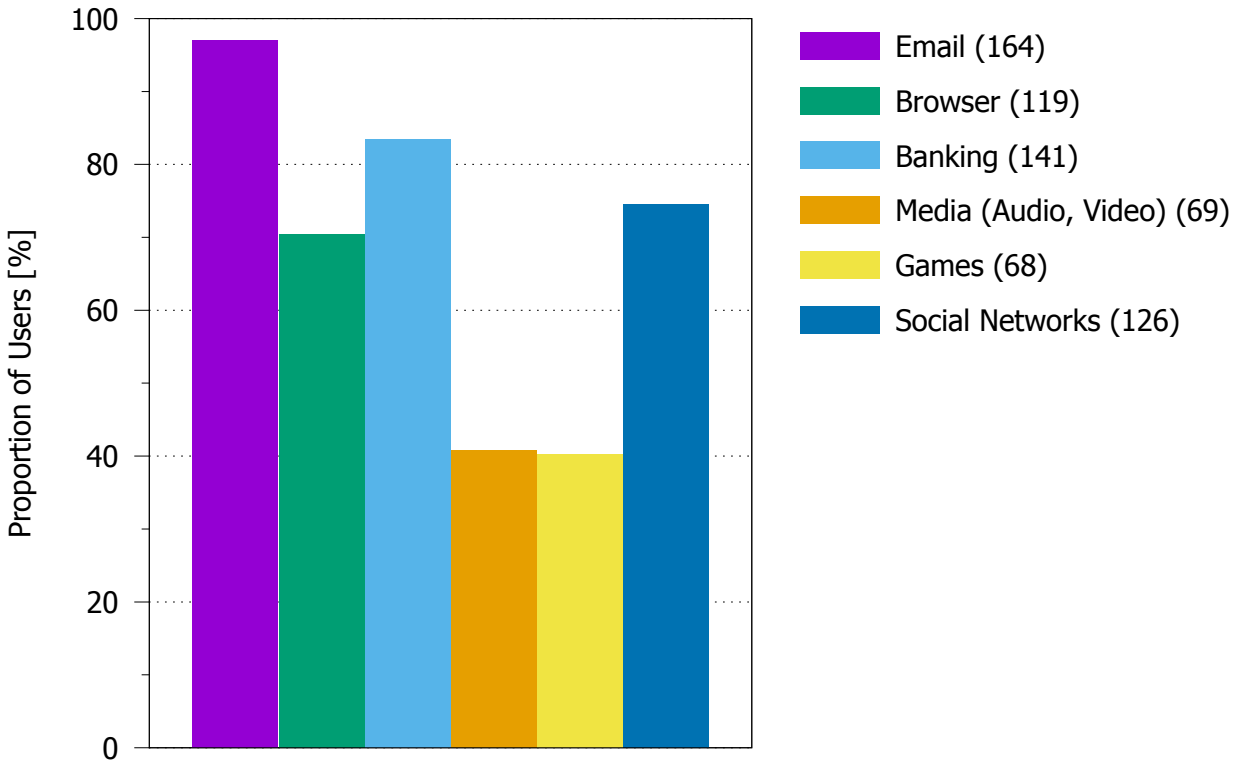


Figure 6: Services Endangered By Phishing

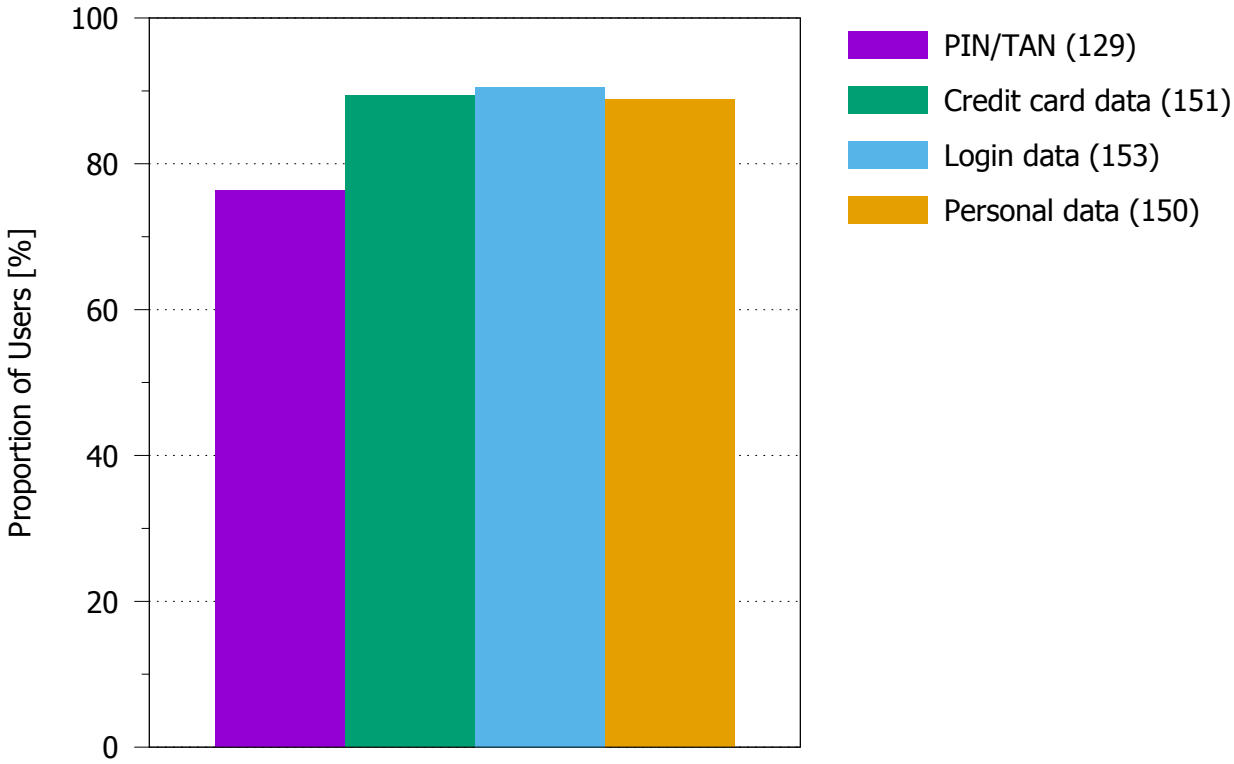


Figure 7: Data Endangered By Phishing

7 Teaching and Learning Content

In this section we describe and elaborate on different teaching and learning contents which can potentially be communicated to the user. At the same time we reason our decision whether to communicate the specific content or not.

7.1 E-Mail Spoofing

The first thing we tell the user in the app is that they cannot trust in e-mails and the sender of them. This is especially important, because, according to a study [26] by DIVSI, 85% of the Internet users utilize e-mails for communication but only 14% have concerns about security regarding it. This is a problem because in contrast to public believe e-mail is in no way secured against fraud. There are three main facts that we need to transport to the user:

From Field The first misbelife is that the from-field is in some way secured. In reality it must be considered a plaintext field. The problem is that most modern e-mail clients hide this fact away from the user. Therefore we show the user with an example that anyone can send e-mail from any from-address.

E-Mail Content We also have to show the user that the content of the e-mail is totally in the control of the sender. Nobody prevents the attacker from sending e-mails that look exactly like the ones sent out by a legitimate sender.

Links in E-Mails The third information that most users are not aware of is that links in general and e-mail links in detail could point to any page. This means that the link does not necessarily lead to the denoted url.

To show the user these facts we have created the first part of the app. We commonly refer to this part as the “awareness part” because it is intended to increase the users’ awareness of the problem of e-mail spoofing and Internet fraud. Thereafter, the user is presented with a form that allows him to enter a sender and a target e-mail address and a free text. When the user submits this form we send out an e-mail including the sender and the receiver that the user entered. The body of the e-mail contains a common introduction and the users free text. It also contains a link that seems to go to a common webpage but instead links back to the app. The template for this e-mail can be found in Appendix A.

7.2 Smartphone Limitations

As already discussed in section 1.4 smartphones have several limitations, such as the small screen size. This section deals with the detection of phishing on the smartphone and the related limitations. More particularly, we will briefly explain in which ways URLs can be checked with the smartphone and what kind of problems these operations raise. Based on this we decided whether to communicate this kind of URL checking to the user or not.

Invisible Address Bar Due to lack of space most of the smartphone browsers hide the address bar [3] and use this expanded space for the web content. By doing this, not only potential security indicators are made invisible, but also the URL that indicates with which website the user is actually interacting. In order to make the address bar re-appear the user generally has to scroll to the top of the whole website. The fact that the address bar containing the important information of the URL is generally hidden must be communicated to the user. Most of the users will probably know that they can access the address bar by scrolling to the top of the website. However, for those who might not know how to deal with that an introduction is inevitable.

Analyze Complete URL Via Address Bar Finding the address bar will not suffice for a reasonable URL analysis. Here again, the small screen size makes it is impossible to view the complete URL without any further action. Specifically, it is necessary to first tap the URL address text and then scroll the pointer to the left and right for the URL analysis. Without learning these steps a reliable URL checking is not possible. Therefore, these steps have to be communicated to the user.

Show URL Before Click - Preview Many mobile e-mail clients provide the functionality of showing the URL a link leads to when touching and holding the link. However, the Android stock e-mail app, for example, does not provide this functionality. This operation is generally available in smartphone browsers for links on websites while surfing. Yet, one should keep in mind that it might happen that the complete URL cannot be displayed on this preview in case it is too

long. Consequently, as discussed in section 4.1, deceiving the user with well-crafted, illegitimate URLs becomes possible. There, we already discussed the benefits and drawbacks of teaching the user how to preview the destination URL and our decision against it.

Show URL Before Click - Copy and Paste Previewing the destination URL raises flaws, such as there is no guarantee that every mobile e-mail client provides this functionality and deception remains possible. An alternative to the preview functionality is the copy and paste functionality. When touching and holding a link the option "copy URL" is available additionally to the URL preview. Upon selecting this option the destination URL is copied to the clipboard. Now the user may paste the copied destination URL to any editor or even the address bar itself in order to analyze it *before* submitting. In case the URL was pasted directly into the address bar, left and right scrolling must further be applied for the analysis. Analyzing the URL in a separate editor would mean to re-paste the URL into the address bar again before submitting it. Since these approaches require several steps before the user may visit the website, we believe that either of those are of too high effort and thus would not be followed by the users. Also, the user would not be able to see the "real" target in case there is a (possibily malicious) redirect included. Hence, workarounds to show the actual url in advance will not be communicated to the user.

7.3 Structure of a URL

The URL is a complex construct. In order to correctly analyze a URL and decide whether it is spoofed or not it is necessary to understand the structure of it. Therefore, the users must achieve the essential skills of parsing a URL appropriately. An attacker uses brands which are familiar to the user in any part of a URL in order to deceive him. Thus, especially the identification of the domain in a given URL is a key aspect which must be covered extensively within the app. We will teach the user about the most important parts of a URL, the protocol (http vs. https), the host with its domain and the path. We do not distinguish the directory, filename or query parameters of the URL path because this part is rather irrelevant for the detection of phishing. By contracting these elements we can avoid discouraging and overwhelming users with details they do not need.

7.4 Phishing URLs

As aforementioned, we focus on teaching the user how to analyze a given URL and to decide whether it belongs to a legitimate or illegitimate website. In order to distinguish legitimate URLs from phishing URLs it is necessary to analyze existent phishing URLs regarding the way they are spoofed in order to deceive the users. For the analysis of phishing URLs we chose the database of PhishTank [67]. PhishTank is a free community site where people can submit, verify and view phishing data. It provides an API which makes all PhishTank data accessible. Organizations such as Yahoo, Kaspersky Lab and Mcafee use the data submitted by PhishTank. A further reason to choose PhishTank as our phishing URL database was that Kaspersky Lab itself recommended to make use of it for our URL analysis. For the phishing URL analysis we drew on the URL categories identified by the authors of Anti-Phishing Phil [79] as our baseline. To these belong IP address URLs, subdomain URLs as well as similar and deceptive domain URLs. Subsequently, we went through the PhishTank URLs and tried to assign them to one of these categories. When no category suited the URL to be assigned, we generated a new category, to which the URL could then be assigned to. In addition we found various categories mentioned in literature, which we also included to our categories, even if we could not find any explicit example URL in the PhishTank database. In the following the identified URL categories are explained.

7.4.1 Phishing URL Categorization

URLs are complex and many users do not know how exactly they have to be interpreted. For example, users can be convinced of the authenticity of URL when it contains the brand name anywhere. Phishers exploit this lack of knowledge in different ways. In the following we present the identified categories of spoofing attacks on URLs. All spoofing attacks are covered by the app unless noted otherwise.

Subdomain Phishers make use of subdomains which are very similar or even identical to the domains of the spoofed target institutions. For example, they register a domain “xyz.com” and use “paypal” in their subdomain, resulting in a URL such as “http://www.paypal.xyz.com/webapps/”. This makes the users believe that they are on a legitimate website.

IP Address Sometimes phishers do not even bother registering any domain at all. In this case, the host area of the URL contains an IP address.

Nonsense Domain We frequently encountered URLs which had registered random names or strings as their domain. The domain names ranged from random letters to domain names like “marketstreetchippy. com”. In order to deceive the users some of these URLs contained a well known brand name in other parts of the URL. Yet, there were a number of examples where this was not the case, i.e. without viewing the content of the website one would have no idea where the URL points to.

Trustworthy, but Unrelated Domain Some URLs are very well-crafted. When reading them they appear meaningful and trustworthy. This is particularly accomplished by registering domain names which sound reliable, for example, “account-information.com”, “secure-login.de”, or “security-update.com”. If the URL additionally contains the brand name of the target institution somewhere in the URL the user is easily deceived.

Similar and Deceptive Domains Another possibility to fool users with a spoofed URL is to use URLs which look like the original ones, but have a slight difference. For example, phishers register domains which resemble the targeted domain, but have a typo. To spoof “paypal.com”, for instance, the attacker might register “paypel.com”. Another approach is to use a modification of the original domain. The modified domain contains the brand name in some form. For example, “facebook-login.com” can be registered in order to fake “facebook.com”. Finally, the attacker can scramble letters of the original domain, which can be very hard to detect at first sight.

Homograph Attack The homograph attack exploits character resemblance. Here characters are replaced by other characters which look very similar to the replaced one. For example, an attacker might replace a “w” within a genuine domain with “vv” and register it. An even more advanced way is to replace characters of the genuine domain with characters from other character sets, such as Cyrillic languages, where the characters will look almost identical [31]. The latter case is indistinguishable for the human eye in many cases and is partially a technical issue. Here, browser vendors should be encouraged to indicate when there are international characters. For this reason only cases that are distinguishable by the human eye are covered by the educational app.

Tiny URLs A tiny URL service is used to convert a long URL into a short one. Due to their shortness tiny URLs are very comfortable to use and easy-to-type. There seemed to be a trend of using tiny URLs for phishing in 2009, in particular in instant messaging services [48]. Tiny URLs usually do not give a hint about the target website and users do not tend to be suspicious about receiving such links from a “friend” what made the use of it for the purpose of phishing quite popular. Tiny URLs redirect the shortened URL to the actual one. As we consider the “analyze URL after-click” scenario for the user education, there is no need of the cloaked URL to be covered by the app.

Cloaked URLs Other phishers integrate an “@” into the URL so that domain names become difficult to understand and the actual destination of a link becomes “cloaked” [2]. For example, the URL http://paypal.com@google.com/ is redirected to http://google.com. As we consider the “analyze URL after-click” scenario for the user education, there is no need of the tiny URL to be covered by the app.

7.4.2 Problems with URLs

There arise two major problems with the detection of phishing attacks based on the URL which are stated below.

1. Some legitimate URLs feature characteristics of phishing URLs.
2. We cannot assure that the users know all website vendors of our URLs (whether legitimate or phish).

The succeeding sections elaborate on these problems and outline how we approach to handle these problems.

Legitimate but Fraudulent Looking URLs We wanted to find out to what extent our determined phishing URL categories from section 7.4.1 apply to authentic websites. For this purpose we browsed the web and looked at the top 50 banks [91] and top 50 online shops in Germany [15]. While surfing on these websites we have recognized that it occasionally happens that a legitimate URL features the characteristic of a phishing URL. This is particularly the case for the category of similar and deceptive domains. There exist sites of vendors which make use of similar domains, instead of never changing the domain and using, for example, subdomains. An example is the website of the Commerzbank. When surfing on Commerzbank's website the domain is "commerzbank.de". As soon as the user is on the online banking part of the website the domain changes to "commerzbanking.de". The same happens on a PayPal website. The regular website features the domain "paypal.com". However, paypal also has a site, where the domain is "paypal-viewpoints.com" By our definition, "commerzbanking.de" and "paypal-viewpoints.com" contain indications of a phishing website. We decided to address this problem by adding a section of final remarks to the app. In this section we tell the user that there might be legitimate domains which are similar to domains they are familiar with and that these not necessarily are phishes. We still strongly recommend the user to directly contact the vendor before submitting data to such websites. The reason why we do not address this problem in the according level is that we do not want to degrade the user's attention by telling him that similar domains might still be legitimate. We do not consider it an issue that the user is told about this later in the app, since it is better to reject a legitimate URL than trusting a phishing URL.

Unknown Website Vendors Despite our efforts of mainly making use of URLs from widely known website vendors there is still the possibility that a user does not know all vendors and thus the respective URL. One approach to address this problem is that the user has to indicate which websites he knows with the aid of a long list of check boxes, for example. We decided against this approach for two reasons: First, if a user only knows a few website vendors, the list of available URLs to be drawn from would be quite short. And thus the game experience will degrade significantly. Second, and more importantly, we cannot expect the users to go through a large list of vendors and let them decide whether they know them or not. This kind of configuration would substantially decrease the usability and acceptance of our app. Currently, we have to let the user learn new vendors. By making mistakes and/or giving correct answers to unknown URLs, i.e. domains, the user will likely gain experience and learn whether a given domain is legitimate or not. For future work one might consider to add a "I don't know" button to the options a user has during a challenge in a level of the app. In this case the user could be explained whether it is a phish or not and why it is so. This option should be punished in some way in order to prevent the user from picking only easy URLs.

7.5 General Recommended Behavior

There are general hints and tips for Internet users which are helpful. The user should be informed about these general recommendations at some point. These aspects are explained in the following.

Data Entry Via Https We should also tell the user that he should only enter data via HTTPS. When the user is entering data via HTTP there are basically two problems. First the user can no longer be sure that he really talks to the person he want's to talk to. This is captured by our precodition that DNS is assumed to be secure. Secondly even if he really talks to the legitimate target site he can not be sure that an attacker is not wiretapping the communication. Therefore the data that is send over plain HTTP can be considered lost. This teaching content will be part of our app, cf. section 8.5.

Do Not Download Attachments Many users download or even open files that they receive via e-mail rather unchecked. This is related to the problem that they trust the from-field of the e-mail. It is crucial to tell them that downloading or even opening a unknown file might infect their system. However, for this work we consider this as out of scope and leave it open for future work.

Data Economy The goal of this app is to prevent that the data of the user is phished. The first step towards this goal is to teach the user to enter his sensitive data as rarely as possible. The idea behind this that websites, including but not limited to phishing websites, might use the users' data in a way that they did not intend. This is considered out of scope and remains as a problem to be targeted in future work.

7.6 Browser Security Indicators

As a matter of fact, there is a major lack of mobile browser security indicators [3, 11]. Yet, there exist some, for example, the padlock for the usage of https. Such signals have the potential to provide relevant information to the user which we would have liked to inform them about. However, besides the lack of such hints there is also the problem of inconsistencies among the mobile as well as desktop browsers. Ultimately, our decision was not to tell anything about these security indicators, as the inconsistencies are too significant even among the standard browser, depending on the device and Android version it is installed on.

Https Padlock All Android standard browsers on various devices we have examined have a padlock on SSL secured pages. Also, one should consider that there are illegitimate as well as legitimate websites where a padlock is part of the web content. Therefore, it is important to teach the users to look for the padlock in the browser, not in the content, to verify that the site they visit is SSL secured, when they enter confidential data. However, some browsers additionally make use of so called favicons, i.e. small website icons. The danger of using such a favicon is that a phisher could use the image of a padlock in order to deceive the user [11]. Moreover, the padlock with/without favicon combinations appear in different ways. While a part of the stock browsers installed on various devices and Android versions we have examined only feature a padlock in case of https websites and no favicon at all, others always display favicons. In the latter case, if https is used the padlock is either positioned right next to the favicon or overlaps it. Due to the variety of possible combinations as well as the deception potential in combination with favicons we decided not to tell the user about the padlock.

Touch Padlock Touching the padlock of an SSL secured website leads to an alert dialog with information about the website. One part of this information is the complete URL of the website the user is currently on. In this case, it would become possible to view and analyze the complete URL without tapping the address bar and scrolling to the left and right. For *some* browsers which additionally display favicons, the above described feature is always applicable. That means, the alert dialog with the complete URL can also be consulted on websites which do not use https. Yet, there are also browsers where neither clicking on a padlock nor on a favicon is possible. Hence, we will stick to our approach, where the user is explained how to analyze the URL directly in the address bar.

Certificate Verification Tapping on the padlock icon results in an alert dialog where the user can select to view the certificate details (“show certificate”). As already stated, the padlock feature is not consistent over the various devices and browser versions. Hence, in these cases a validation of the certificate is not possible as well. Therefore, this is an aspect which is not covered by our app.

7.7 Conclusion

This section briefly lists the above described learning contents which will be addressed by our app.

1. E-mail spoofing
    a) Do not trust the sender
    b) Do not trust the content
    c) Target URL of a link is not necessarily the displayed one
2. Invisible address bar
    a) Access address bar
    b) View complete URL
3. Structure of a URL
4. Phishing URL categories

a) Subdomain attack
b) IP address attack
c) Nonsense domain
d) Trustworthy sounding, but unrelated domain
e) Similar and deceptive domain
f) Homograph attack

5. General recommended behavior
    a) Data entry via HTTPS only

8 Approach for Our Anti-Phishing Education App

This chapter presents our final approach for the app. In the following sections we will elaborate on the app design in detail.

8.1 App Design

We have decided to divide our education app into two main parts. The first part of the education app is intended to increase the user awareness. The second part of the app then covers the actual educational part. The following enumeration summarizes the functions of our twofold app structure.

1. *Awareness Part* The first part of the education app is intended to increase the user awareness regarding how easy it is to spoof e-mails and mislead users with such fake messages. This part is supposed to motivate the user to do something to counter the danger of the Internet and phishing, in particular.

    a) *Receive Fake E-Mail* We want to illustrate to the user how easy it is to spoof the "from" field of an e-mail as well as the content of this e-mail. For this purpose the user has to send a fake e-mail with an arbitrary sender address of his choice to himself. The user will also have the option to type in a text of his choice. Upon submitting the form the user will receive an e-mail with the e-mail address he had indicated as sender. The entered text will also be part of the received e-mail. We believe that the user will be very surprised about how easy even he himself could send a fake e-mail. Hence, the user learns that he cannot fully trust the "from" field and the content of the e-mails he is receiving.

    b) *Link Text Unequal Target URL* The awareness part of the app is also supposed to show the user that he cannot trust the texts of a link he is clicking on. To illustrate this, the user is asked to click on a link with the text "https://www. google.de/". Clicking on this link, the user will expect to land on the Google website, what will not happen. In fact, the user is linked back to our app, where he is told that link texts are not trustful as well.

    c) *Fake Website* Finally, the user is told that creating a copy of a website is also very easy. He is told that a reliable way to decide whether a website is a fake or not is to analyze the URL of the website he is visiting. Ultimately, he is told that this will be the focus of the following exercises.

2. *Educational Part* The second part of the app covers the actual educational part. Here, the user is learning about various spoofing techniques of the attacker.

    a) *Learning Part* The second part of the app is divided into levels of increasing difficulty. The user is first taught how to access and analyze the URL of the web browser. Subsequently, the user learns about the general structure of a URL. We tried to explain these facts as simply as possible to allow even unexperienced users to follow. In particular, the user is told how to find the second- and top-level domain of a URL. In all succeeding levels the user is introduced to various URL spoofing techniques of a phisher. The learning content of each level is described in section 8.5.

    b) *Exercise* After every introductory material on each level, an exercise section follows. For the "access and analyze URL part", for example, the user is forwarded to a website. There he has to apply all important steps he has learnt in the introductory part. After successful completion of the tasks, the user is linked back to the app. For the "find domain" information material the user gets a couple of valid URLs of which he has to identify the domain. All subsequent level exercises are structured as follows: the user is presented a URL. He has to decide whether the presented URL is a phish or a valid URL. If the URL is a phish, and the user has correctly identified the phish, the user has to show the domain.

    c) *Increasing Difficulty* There is an increase of difficulty for each succeeding level. That is to say, in each level it gets more difficult to distinguish phishing URLs from valid ones, cf. section 8.5.

In the succeeding section we describes common elements in games that we adapted for our app.

8.2 Gamification

As we pointed out we decided on designing the app as a game. This is mainly because we think that this makes it more appealing to the users. In most modern games we see the following elements which we all implemented:

Lives An often used game element are lives. An inherent property of a game is the possibility of loosing the game. If you are not able to loose a game you will get no positive feedback from winning the game. On the other hand, you do not want your player to loose the whole game when he makes one little mistake. Therefore, most games have some kind of "you have N tries"-element, which is commonly referred to as "lives". We also included such a mechanism in our app. The details are layed out in section 8.3.

Levels Also most games have some kind of level system. This has multiple purposes. First it is important for the player to get a feeling for the process he makes in the game. It also provides the user with fixed points in the game where he can restart or pause and play on later. To us it gives the ability to structure the game in a repeating cycle that the user learns. This way if the user plays multiple levels in a row he can skip the repeat parts easily but if he has paused he can read it. The details of the leveling Strategy can be found in section 8.4

Leaderboards Finally often games have a kind of leaderboard. This means a area where you can compare your progress within the game with others of the game. For many people this is a major motivation point. These people want to be better then others. To motivate these people we introduce two leaderboards:

1. Total Points First there is a leaderboard that shows how many points you gained while playing the level. The details of how the points are calculated are shown in section 8.4.
2. Detected Phishing URLs As the points are a relatively opaque number we also introduced another leaderboard. This shows the user that detected most phishes in the app.

Achievements There is another type of player. This player wants to find everything in a game and is willing to invest time in a level just to finish it perfectly or to find every hidden secret. To address this type of player modern games often something called "achivements". Achievements are special parts of a game that you can unlock if you for example find a special object or play a given level very good. We implemented mainly achivements for finding 5,10,25,50,100 and 500 phishing URLs.

8.3 Game Rules

The educational part of the app which is followed by the awareness part is divided into several levels. In each level the user is provided with specific informational material. After the information material is consulted by the user, he has to finish the according exercise. The first and second information materials (introduction 2 and level 1) and exercises that the user receives differ from the ones of the other levels. Here, the users obtain basic knowledge in order to bring them to the same knowledge level for the actual game.

Basic Knowledge The first information material and task of the user deals with accessing the address bar of a webbrowser and viewing its URL completetely. To prove that the user has understood how to access the address bar and view the URL he has to do the following: When the user is forwarded to the website to solve the task he has to scroll up to the top of the website to make the generally hidden address bar visible. Afterwards, he has to provide us the information we request about the URL in the address bar. This will show that he has in fact viewed the whole URL. After successful completion the user is linked back to the app and level 1 is started. From this level on, the user has three lives upon start of each level. In level 1 he has to identify the "Who-Section" (second- and top-level domain) of a URL. He has to tap the according part of the displayed URL. In this level wrong answers result in losing points and losing a live. In order not to frustrate the user he cannot get less than 0 points. When the user has no more lives left he has to restart the level. With every correct answer the user gains points.

Actual Game In level 2 we start introducing URL spoofing techniques and the user has to decide whether a given URL is a phishing URL or a valid one. Here also, the user can lose and win points as well as lose lives. Here again, if the user has no more lives left he has to restart the level. Figure 8 illustrates the game flow and consequences of wrong and correct answers from level 2 and upwards. If the user has correctly decided that a phishing URL is a phish, he has to show us the "Who-Section" to prove that he has understood the concept. In all other cases the user is directly shown the result of his answer. In summary, the user looses points for any wrong answer, but he does not lose a life for every wrong answer. We have decided that rejecting valid URLs is not as severe as accepting phishing URLs. For this reason the punishment for accepting a phishing URL is more severe than the punishment for rejecting a valid URL. All in all, the user loses points and a live in the following cases: the user has falsely accepted a phishing URL or the user has correctly rejected a phishing URL, but could not show us the "Who-Section". In all other cases the user cannot lose lives, but only points.

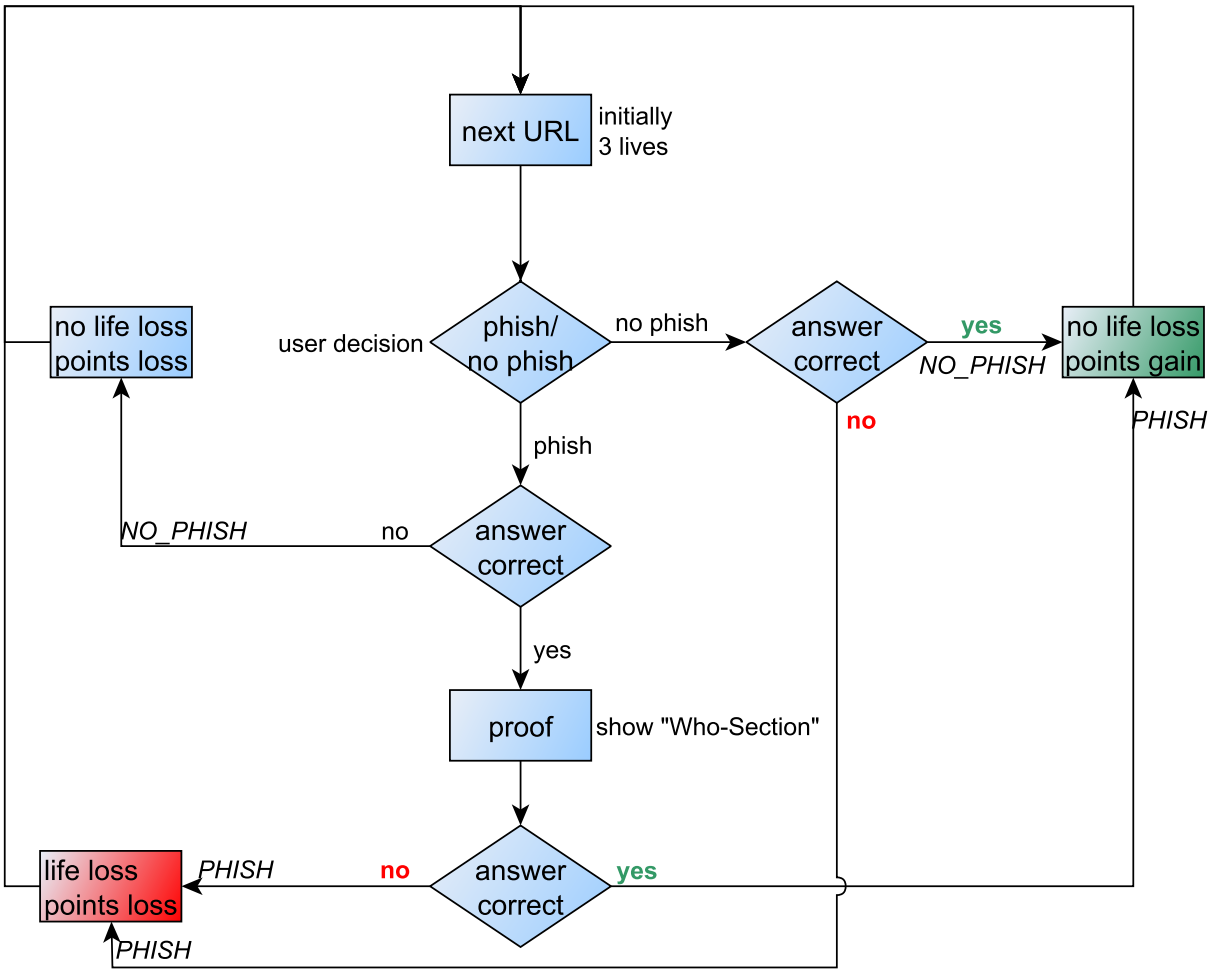


Figure 8: Losing points and lives in the game

8.4 Leveling Strategy

During the app development we have tried out several leveling strategies. This section is intended to introduce the leveling strategies we have considered for the app.

Leveling Based on Achieved Points Our very first leveling strategy was based on the achieved points per level. Each level the user had to achieve at least 100 points to pass the current level and unlock the next one. This approach had a major

drawback. The fact that achieving a minimum of points to pass the level resulted in very similar points for everybody finishing a level. That is to say, everybody who has finished level x, has approximately the same points, which in turn would have meant that the comparison between single users would not be meaningful as it would only differ very slightly. Additionally, with this strategy, users might replay early levels which are easier and gain the same amount of points as users playing later levels. This might result in users playing early levels repeatedly get more points than users playing later and difficult levels.

Leveling Based on Detected Phishes The previously described leveling strategy had the deficit of comparability among the users. However, we consider comparability very important since it serves as an incentive for the user to play better or play on. For this reason we overthought our strategy and decided that passing a level should not depend on the points a user receives. It rather should depend on the number of phishes the user was able to detect during a level. That is to say, among the shown URLs in every level there is a certain amount of phishes the user has to detect in order to pass the level. With this approach however, there is still the possibility for a user to repeat early, and thus easy, levels and possibly gain more points than users playing later, and thus more difficult, levels. To prohibit this, in increasing levels the users gains and loses increasing points accordingly. The points per correct answer increase such that it does not pay of to play lower levels. The points for each answer p in each level n is mudified by the formular:

$$p * 1.5^n$$

In this way, a user repeating early levels is not able to catch up other users of higher levels. This strategy solved the problems of our first strategy, however it also brought a new one. The strategy of passing the level when a certain amount of phishes are detected has the following flaw: always rejecting a URL will eventually result in passing the level (if the user also correctly identifies the “Who-Section” when required). The user will not gain a lot of points with this strategy, however he will eventually win, which is suboptimal for a game.

Leveling Based on Correct Answers To solve this problem of our second leveling approach we have extended the leveling passing to correct answers. Instead of detecting a certain amount of phishes per level, the user has to give correct answers to a predefined amount of phishing URLs as well as a predefined amount of valid URLs in order to pass the level. Only and only if the user has answered the predefined number of valid and phishing URLs the level is completed. To additionally incentivize the users we have included three lives per level. The lives are supposed to prevent a user playing eternally, without ever passing the current level. When the user loses all of his lives, cf. Figure 8, this is an indication that he did not understand what the level is about. Consequently, he has to restart the level by being forwarded to the introductory part of the current level. The points are assigned exactly as before. This is our final leveling strategy for the app.

8.5 Teaching Goals per Level

This section summarizes the learning objectives of each level. Note that we generally do not use technical terms like URL, domain, subdomain, protocol or the like. Figure 10 illustrates and exemplifies the level flow of our app.

Introduction 1 This part is the awareness part described in section 8.1. Here, the user learns how easy e-mail spoofing is. Additionally, the user is informed about the simplicity of setting up fake websites and that he should not trust the texts of the links he is clicking on.

Introduction 2 In this part the user is explained how he can access the URL of a web browser and how exactly he has to look at the whole URL. In particular, the user is told that he has to scroll up the whole website to make the generally hidden address bar re-appear. Then he has to tap the text field of the address bar and scroll to the start of the URL. At the end of the exercise for this the user is told that he always has to analyze the URL like this, because all other displayed URLs or links might be fake too.

Level 1 The actual game starts with level 1, where the user learns about the structure of a URL. First of all, the user gets an overview of the single components of a URL. To make the comprehension of these components easier to understand

we used an analogy which is summarized in Figure 9 with an example URL. We told the user that he has to imagine that the website he is visiting is his dialog partner. The user is told that the section between "http(s)://" and the third slash "/", i. e. the hostname, reveals information about his dialog partner. In particular, we explain that he has to read this part from right to left. The top-level and second-level domain is introduced as "Who-Section" (company + location of company), from which the user knows who he is actually talking to. All succeeding parts in this area are to be considered as "departments" of the company of ther user's dialog partner. The protocol part is introduced as "Security Level" of the dialog with the partner and the path part of a URL, i. e. the part after the third slash "/", is introduced as the topic of the conversation with the dialog partner. When marking parts of a URL we consistently used the according colour of Figure 9. The main objective of the level 1 exercise is to be able to identify the second- and top-level domain of a URL.

Level 2 With level two we start introducing the spoofing tricks of a phisher. We considered the subdomain attack, cf. section 7.4.1, as a good starting point to introduce the phisher as the user has just learnt about the importance of the "Who-Section" (top-level and second-level domain) in level 1.

Level 3 In level 3 the user is first told what an IP address is. To facilitate the comprehensibility, we used the analogy of house addresses. The user is explained that like addressing our houses with street names and numbers, computers in the Internet are addressed by so called IP addresses. The IP address itself is defined as a 4-place sequence of numbers, separated by dots. Finally, the user is warned against URLs with IP addresses in the host part.

Level 4 In this level we deal with nonsense in the second-level domain, cf. section 7.4.1.

Level 5 In this level we deal with second-level domain names which sound trustworthy, but are in fact unrelated to the company name, cf. section 7.4.1.

Level 6 Here misleading and deceiving names in the second-level domain of a URL are covered. This includes typos, scrambled letters or other similar and deceptive names in the second-level domain, cf. section 7.4.1.

Level 7 In this level we focus on homographic attacks, where the user is able to visually distinguish a fake second-level domain from the original one, cf. section 7.4.1.

Level 8 In this level the user is introduced to an attack where the brand name of the visited website or even the whole legitimate URL is placed in the path of a fake URL, cf. section 7.4.1.

Level 9 Here we introduce the difference between the usage of http:// and https://. In particular, the user is told that the usage of https:// means that his conversation with the website is encrypted and that the dialog partner indicated in the "Who-Section" is authenticated. As an analogy we say that the https:// represents a higher security level. This means, the conversation cannot be eavesdroppbed by a third party and the dialog partner indictated in the "Who-Section" has proved his identity to a trusted third party. With http:// this security level is not established.

Level 10 This level does not include an exercise. It mainly serves as a section with some important additional input for the user. Specifically, we tell the user two things: First, we explain to him that he might encounter URLs which actually look very phishy. In such a case, we suggest him to directly contact the company and ask for the authenticity of the specific website. Furthermore, we introduce extended validation certificates. We provide the user with a link to further information to this subject.

8.6 Use of Learning Principles and Game Techniques

Our app is a learning game which purposes to introduce information on the topic of phishing and how to detect it. With the app we want to improve understanding for this topic and help users be less vulnerable for falling for such attacks in future. This section deals with the principles of learning and game techniques which are reflected in our app design. In fact, the laws of learning and game techniques have a strong connection which is the reason why games work for learning purposes [59].

Figure 9: URL components that are communicated to the user

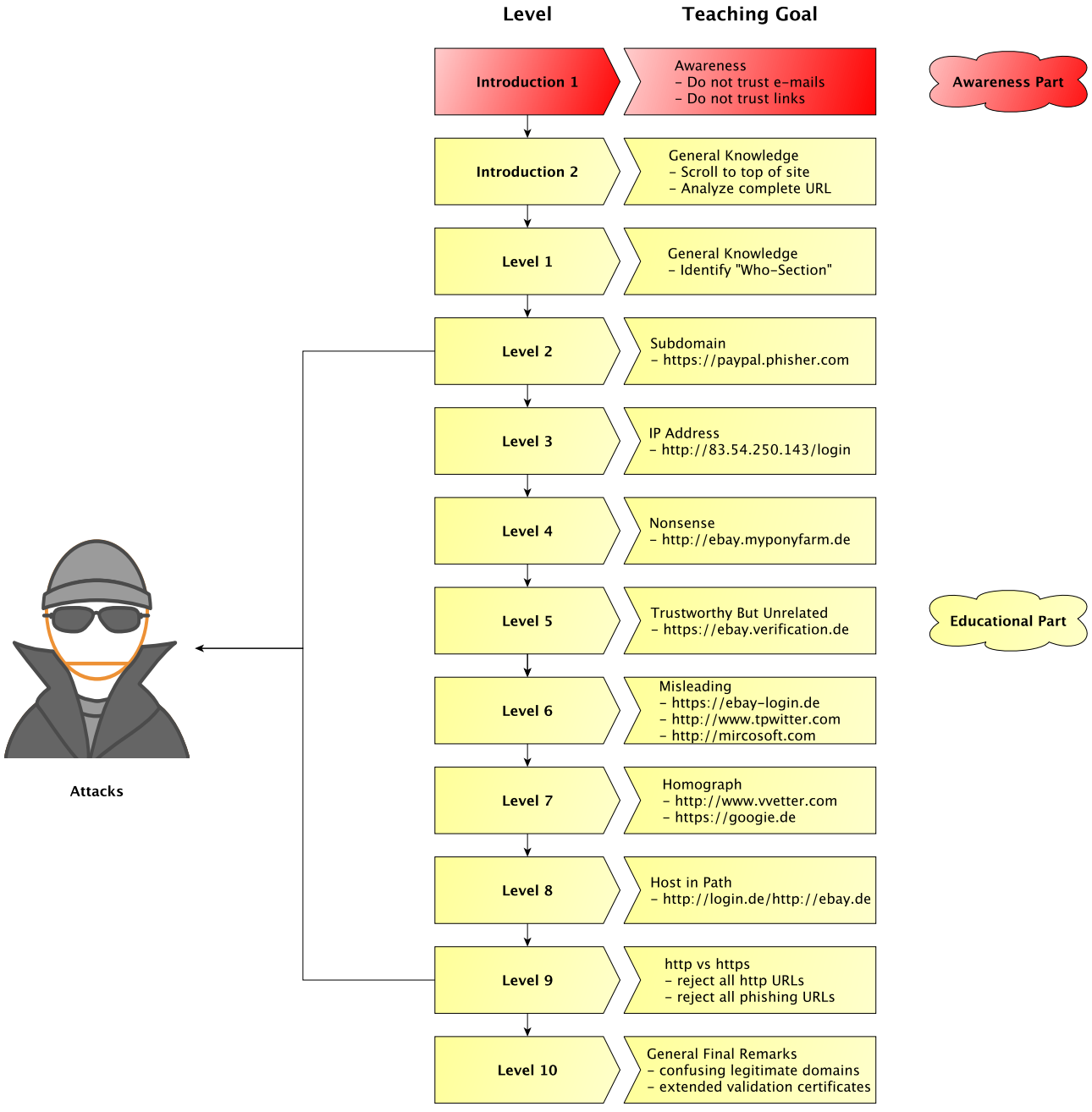


Figure 10: Teaching goals of the app

8.6.1 Principles of Learning

Edward Thorndike introduced the first three principles of learning: readiness, exercise and effect [87, 59, 39]. Later the principles of learning were further extended by: primacy, intensity and recency [59, 39]. These principles outline how people learn and which conditions improve their process of learning. In the following we will briefly introduce the meaning of these principles [59] and explain how they are reflected in our app.

Readiness The principle of readiness claims that physical, mental as well as emotional preparedness is an important prerequisite for a better learning performance. It also states that motivation is crucial for effective learning. First of all, students must want to learn something, otherwise any additional motivational efforts will be of no use. In order to make students want to learn something it is relevant for them to see a clear reason for learning, i.e. the perceived value of the material is ultimately related to their motivation. Finally, the principle of readiness says that best learning performances are achieved in combination with good physical health. The physical and health conditions of our app users are beyond our control. For this reason this is an aspect which cannot be reflected in our app. The app is targeted at users who are willing to learn something about phishing. Consequently, there is already some kind of motivation in our app users. In order to increase this motivation we present clear reasons why the user should continue to play our app. This is happening in the awareness part, where the user is told what exactly phishing is, how easy it is to phish users, spoof e-mail senders and content, spoof links as well as create exact copies of legitimate websites, cf. section 8.1.

Exercise The use of exercise is composed of two important parts: First, training and repetition help increase learning. Second, feedback is crucial for good learning performance. For best learning results, these two parts must be applied together. In a nutshell, the learning connection is strengthened by practice, and weakend by disuse [39]. After finishing the awareness part, we start with the user education and provide exercises for all of our learning goals (except for the last level), cf. section 8.5. The learning content is also permanently repeated. For example, in the introduction 2, cf. section 8.5 the user learns to scroll up to make the address bar re-appear and analyze the URL by right and left scrolling it. The scrolling of the URL is present in levels 2-9. Also, the user has to identify the Who-Section in level 1. In addition, the user has to show the Who-Section every time he has detected a phishing URL correctly. Finally, every introduced attack of each level n will appear in level n+1 at least once. That is to say, as soon as the user gets to know a new attack, he will keep seeing this attack in the succeeding levels. Our app also provides direct feedback to the users' actions. In case of correct answers, the user is rewarded with gained points, with a smiley and a small text that he has done well. When the user has made a mistake he is punished with losing points, possible life loss and a sad smiley. Also, told that the answer was wrong, why it was wrong and additionally he gets a reminder text on the applied URL spoofing attack he was not able to recognize. To sum it up, we let our app users practice what we taught them and make use of repetitions in combination with direct feedback.

Effect A student who associates his learning with positive feelings will learn more and better than another student who connects his learning with negative feelings. For example, a student who is unsuccessful with initial learning material will associate his experience with unpleasantness, frustration, anger and/or confusion, while a student with early success will have strong positive feelings and thus will be more motivated to have more success in future. Therefore, enabling particularly early success and maintaining student motivation with positive feedback and comments is crucial. In our app early success is easy to achieve, since we start with easy tasks and obvious attacks which get increasingly difficult in higher levels. When the user has given a correct answer he is rewarded with a big smiley face as well as points. Also, the user is shown a medal every time he finishes a level. These aspects of the app are intended to increase the users' positive feelings and keep him motivated to go on. However, some improvement of positive comments might be worth considering. Currently, the positive texts for a correctly answered text, for finishing a level and icons do not differ. It might be more engaging if such texts and icons differ from time to time in order to increase the positive emotions of the user. For example, texts telling the user that he has done well might slightly vary (for instance, according to the degree of difficulty of the achieved task). Also, the screen of finishing a level could vary. Here again the degree of difficulty might be considered. According to the level finished, the obtained award could get bigger, the text of the level finished screen more flattering in order to increase the users positive experience.

Intensity This principle says that learning is better encouraged by things that are more intense. For example, people likely to learn more from an exciting and enthusiastic teacher than from a boring and monotone one or from a text book. Our app is by nature more intense than a simple text based approach. The game creates incentives and intensity. Also, the fact that we do not only tell the user that e-mail spoofing and link spoofing is easy but also make him experience it increases the intensity.

Primacy The principle of primacy means that the first thing a student learns makes the strongest impression. For this reason getting rid of bad habits, and replacing incorrect or wrong logic are difficult. This principle is coupled with time. The first things our users learn is: what is phishing as well as how easy e-mail and link spoofing is. For those who already knew these things, the awareness part will not be such a high motivating factor. Yet, we believe this part is indispensable since there are still many people outside who are not aware of the aspects mentioned above.

Recency The principle of recency states that most recently learnt things are easier to remember. This is a consequence of the reduction of the learning over time. The principle of recency is coupled with time. We are aware that our app is not a game which will be frequently used and thus its users are likely to forget the content they have learnt, for example, a week ago. In order to overcome this problem, so called reminders are included in each new level introduction. There the user has a short summary of what he has learnt so far. Also, we keep confronting the user with attacks from previous levels during the exercise rounds. This repetition is, on one hand, intended to strengthen the users knowledge and understanding, on the other hand, it is intended to create a kind of recency so the user does not forget about this kind of attack. In case the user does not detect the attack (a repeated or a new one) he will be reminded what kind of attack had been applied. Furthermore, in this level the user will be confronted with this kind of attack again, until he gives a correct answer to it.

Now that we have introduced the fundamental principles of learning and associated these with our app, we proceed with the introduction of game techniques and how they are related to the learning principles as well as to our app. As already mentioned, there are strong connections between learning principles and game techniques. Therefore, we will try not to focus on redundant aspects, but rather on additional aspects which need to be considered in game design.

8.6.2 Game Techniques

Basic game techniques are: flow, feedback, simplicity, immersion and engagement, choice and involvement, practice as well as fun [59]. As some terms already reveal, these principles are strongly connected to the basic principles of learning. In the following we will elaborate on these game techniques by stating their relation to the according learning principle, mentioning additional aspects to consider and how these are mirrored in our app.

Flow Flow is the key point of games. It is “the state in which people are so involved in an activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it even at great cost, for the sheer sake of doing it” [22]. Sometimes flow is also referred to as ’engagement’ [59] and relates to a person’s overall well-being [78]. Flow relates to motivation [22, 23]. Motivation in turn is a crucial part of readiness, cf. section 8.6.1. In essence, there are four requirements for flow [22, 23, 74].

1. *Clear Tasks* With clear tasks the user is able to understand what he needs to do. The tasks which need to be completed by the users of our app are never complex and they are always clearly told what they need to do next.
2. *Feedback* With feedback the user should always be kept up-to-date about his progression towards the goals he is asked to achieve. He also should get immediate feedback on whether his actions are good or not. Our app covers these aspects, cf. section 8.6.1.
3. *Balanced, Attainable Goals* The user should be confronted with challenging tasks, but at the same time these tasks should also be achievable. Especially in the beginning our app users are confronted with very simple tasks. For some users they even might be too easy which may result in a loss of interest. However, these tasks are important basics which are necessary for successful detection of phishing attacks on the smartphone. Therefore, for future work especially the first two tasks (access address bar and analyze the complete URL) could be re-designed so that

they also keep users which have already knowledge in this area. Currently, the users can just skip the introductory part of this part and directly complete the task. Besides, as the users' skills will naturally improve, their tasks get more difficult and challenging with increasing levels, but will remain achievable.

4. *Concentration* The user should not be distracted with, for example, complex interfaces. He should rather be able to fully concentrate on the game. Our app has a very simple user interface with the most necessary elements. There are no special effects, advertisement or other elements which might distract the user from playing the game with full concentration. The only intrusive and interruptive elements are our introductory sections. However, these are inevitable for the communication of the learning content.

Feedback Feedback is important which is also reflected by the fact that it is a crucial part of the learning principle 'exercise', cf. section 8.6.1, as well as a requirement for flow. Stated simply, feedback is how a user perceives progress [22, 23]. For the completion of even simple tasks feedback is indispensable. Feedback can be in form of a scoring system, comparative statistics or failure outcomes and provides the user information about his progression and performance. Games make use of a so called feedback loop [32]:

1. *Measure Behavior* Our app assesses whether the answer of the user to a given task is correct.
2. *Relay Measurement to User* The user is told whether his answer is correct or not.
3. *Realize Some Sort of Outcome* The outcome of the users' actions and answers are accordingly defined, cf. section 8.3. For example, the user loses a life in case he did not detect a phishing URL.
4. *Provide Opportunities for Alternate Action* The user has the chance to do better in the next tasks.

Simplicity The real world is a complex construct. However, games should simplify the real world so that there only remain rules and goals. In this way the players can fully concentrate on their tasks and how they can achieve them [23]. Hence, simplicity helps users to achieve flow and thus increased motivation. This, in turn, leads to improved learning, cf. section 8.6.1. Simplicity involves, for example, the user interface, the game goals, feedback loops, rules and instructions. The structure of our app is kept simple and consistent. Therefore, it should be easy to understand. Our user interface is kept to the necessary minimum and the goal of the game is clear: detect phishing URLs.

Immersion and Engagement Immersion involves a passive activity. The term is used, for example, to describe a person who shows strong interest for a story [54]. In contrast to immersion, engagement involves active actions, such as trying to solve a problem or puzzle. Games commonly use both, immersion and engagement. To achieve immersion game designers make use of stories, visual and audio techniques, attractive graphics or animations [74]. Simultaneously, the user is engaged with choices, problems, or puzzles which have to be solved. The combination of immersion and engagement has the potential of creating an intense game experience [59]. These two aspects of game techniques are strongly linked to the learning principle of intensity. By challenging the user with various tasks to solve we meet the requirement for achieving engagement. However, immersion is an aspect we have not considered yet in the scope of this thesis. In order to achieve an intense game experience, this aspect might be worth considering for future work.

Choice and Involvement Games consist of choices and involvement. There is a link between choice and positive feelings (cf. Principle of Effect in section 8.6.1), i.e. choice is important for a person's overall well-being [78, 76]. However, the downside of choices is the so called paradox of choice which states that choice is beneficial, but too many choices can cause more bad than good [76]. The problem is, when the users are confronted with too many choices they get overwhelmed, since the decision thus the task to solve becomes too complex. We believe that our app does not face the problem of this paradox since the decisions the user has to take are limited to the questions: is the following URL a phish, and show us the Who-Section. Still, the user has to make decisions and is consequently involved in the game.

Practice This technique is directly related to the learning principle of exercise, cf. section 8.6.1. Users practice and repeat several steps of games extensively so that they eventually gain mastery and the difficulty of their challenges can increaseč̌itemurphy2011games,schell2008art. Our app offers practices as well as repetition, cf. section 8.6.1.

Fun Fun is an important aspect of game design and yet the definition of it is not clearcut in literature [59, 74, 45]. Based on several definitions found in literature Curtiss Murphy introduced the following definition of fun: *"Fun is the positive feelings that occur before, during, and after a compelling flow experience"* [59]. Positive feelings include, but are not limited to, engagement, enjoyment, pleasure, entertainment, satisfaction, control and triumph. Fun is related to the learning principle of effect and its positive feelings. How we achieve the principle of effect and positive feelings in our app is described in section 8.6.1. Yet, fun is something which emerges from several game techniques, such as, flow, immersion and engagement, practice to achieve mastery, and choices, which all lead to positive emotions. Fun is an aspect of our app which could be considered more deeply in future work. Especially, the areas of creating positive feelings and including immersion in order to make the users' game experience more intense and fun are aspects which might be looked at.

9 Development Process

This chapter deals with the development process of our app. We do not provide in-depth insight to our source code. Instead we give a brief overview of our approach for the development of a user friendly and understandable app.

9.1 Mock Up

After we have decided about the work flow and structure of our app we built a mock up in order to get a more concrete idea of what needs to be implemented and to reveal flaws in our thought process. Also, we showed it to a couple of friends and relatives so we could expose aspects we have not yet thought about. All in all, the work flow and structure of the mock up was quite understandable. However, the first texts explaining how to access the address bar and about the structure of a URL seemed to be incomprehensible. As a consequence, we adjusted these texts in the app (only thos of the first three levels) and showed them to other friends and relatives who seemed to understand the descriptions. Based on these initial texts we wrote all remaining texts without including it into the app yet. The next section deals with the elaboration of these texts.

9.2 Pilot Study of App Texts

The app texts were written down in a Google Docs document. After finishing the texts for each step of the app flow our supervisor, a professor of pedagogy at TU Darmstadt as well as another schoolteacher reviewed our texts and gave their feedback to it. As we achieved the version with which we were satisfied we applied a small user study on the created texts. For time reasons we decided to go for the low cost method of guerilla user testing [34, 80]. This approach enables to quickly assess the effectivity of a design, in our case our app texts. Guerilla user tests are rather loosely structured and do not include participant recruitment. The testers are rather approached, in our case, we approached relatives and friends. The outcome of such studies are rather qualitative, i.e. extensive and detailed insights is achieved. A downside of guerilla testing is that the approached participants might not belong to the defined target group with respect to their expertise or skills. Since we knew our participants we are confident that they matched target audience. In detail, our approach for the guerilla user test was as follows:

1. *Prepare Texts* Our aim for this user test was to imitate the use of a smartphone as best as possible. For this reason the app texts in the Google document were formatted into short lines, so that the text appearance resembled that of a smartphone screen. Furthermore, we printed out the texts and cut the sheets into small rectangles.

2. *Think Aloud* We asked the participants to think aloud during the test. We told them that there are no stupid questions or comments and that they help most with just saying what goes through their mind. We made notes of their remarks.

3. *User Test with In-Between Exercises* The actual user test mainly consisted of reading our app texts and thinking aloud about these. We included a little simulation of our exercise parts in order to validate whether the users comprehended the texts or not. For example, for each introduced attack we included a small list of URLs on which the users had to decide whether they were phishing URLs or not.

4. *Final Comments* After going through the texts the users were asked to give general feedback about their impression of the texts. We further asked them about some aspects we were not quite sure about at the beginning. For example, we asked them whether the usage of the terms link or web address confused them.

Our guerilla user tests showed that our texts are understandable. According to our participants the main downside of the texts was their length. Yet, this can be neglected since the users had to read our complete texts (instead of for example just playing 1-2 levels at once). Furthermore, they remarked that the introduction on how to access the whole address bar and analyze the complete URL is unnecessary. For some users this might apply. However, it is possible that there are users who do not know this. For those, who already know how to access the address bar and analyze the complete URL we added a button which directly links to the exercise. In case the user had overestimated himself, he will

be forwarded back to the app, where the introductory text can be consulted. Finally, the reminder texts received some criticism for their frequent re-appearance at the beginning of each level. This can also be neglected since we assume that our app users will not constantly play this game. Also, when playing the app this screen can easily be skipped as exhibits a recognition value achieved by the title “reminder”. Still, we decided for a minor reorganization of the reminder view. Before the user tests the reminders mainly referred to the URL structuring they have learnt so far. We thought it is also important to remind the users of possible attacks. Therefore, the reminder concerning the URL structure was kept to a minimum with the aid of a graphic. Additionally, for each attack in previous levels one sentence and one example was added.

9.3 Implementation and Testing

In parallel to formulating and testing our app texts we developed the basic structure and logic of our app. After conducting and assessing the guerilla user tests with our texts and integrating the feedback we started to merge the texts with our app. We developed and tested the app simultanously. Occasionally, we showed the app to friends and relatives in order to get some feedback on aspects we might have missed. In this way, our app was formed incrementally. We do not intend to provide in-depth implementation details in this work as there is no need for this. The only implementation detail that might be of interest is how our app generates URLs on which the users have to decide whether they are phishing URLs or not. Our URL generation approach can be consulted in Appendix B.

10 Evaluation

As a final step the app we have designed and implemented needs to be evaluated which is the goal of this chapter. The app will be evaluated with the aid of a user study. After introducing our study design, we will state our hypotheses and explain how we are going to measure our statements. Finally, we will analyze our results and state our conclusion.

10.1 Participant Recruitment

This section deals with the participant recruitment for the study. In order to motivate users to attend our user study a 20€Amazon gift was raffled among 4 users. To reach potential participants we proceeded as follows:

Flyer We prepared a flyer with the most important information. In this flyer we told the user that a learning app about Internet security in general will be tested. We did not mention the specific topic of phishing because we did not want potential participants to read up about it in advance. One might question our approach of not telling the user what exactly the study is about. However, we do not consider this an issue since the user was told about the exact topic at the beginning of the study. Copies of this flyer were hung on blackboards of student dorms and some other buildings at the university.

E-Mail to Professors We additionally distributed the flyer to a number of professors in our university and asked them to forward it to their students and/or teaching stuff. We did not forward the flyer to computer science professors or the like since their students and stuff most likely does not match our target group.

Online Social Networks We contacted our friends in online social networks and asked them to ask friends whether they would be willing to participate in our user study. Additionally, we posted the flyer in university groups of online social networks with the hope some people might be interested in participating.

Further Networks Finally, we called friends we could not reach via online social networks and asked them if they knew anybody who would participate. Also, we tried to approach random people at the university, however this was rather unsuccessful.

In section 10.1 copies of our e-mails to the professors and of our flyer can be consulted. Note that we had to split the user study into groups of 4 participants each due to the lack of available smartphones. Moreover, our flyer originally said that the best participant of each group would win the gift certificate. However, we recognized that the winning chance might not be equal for every participant, for example, due to varying expertise or other possible technical problems, such as app crashes. For this reason we asked the participants at the beginnig of each study whether it was okay for them to raffle the gift certificate instead. Here again it is ethically questionable to award only one of the participants with a gift certificate while others might come away empty-handed. To address this problem we decided to offer cookies and other kinds of sweets. Additionally, the participant who performed best was awarded with a “Golden Anti-Phish Certificate”, all other participants received a “Silver Anti-Phish Certificate”.

10.2 Study Design

For our user study we chose a within-subject design, i.e. a “before and after app” study with the same group of people. The advantages of this design can be summarized as follows: Within-subject deals better with variability associated with individual differences compared to between-group design, where different groups would be considered who do and do not play the app. Furthermore, if we had decided for the between-group design we would have needed twice as many participants which we would have not been able to recruit. A major drawback of the withing-subject design, however, is the learning effect. We are not able to clearly distinguish whether a behavior change after playing the app is a result of the app intervention or whether the paricipants just have overthought their decisions on questions they had replied to before. Yet, our results showed that this learning effect can more or less be ignored (cf. section 10.5.3). Figure 11 illustrates the structure of our study, which we explain in the following:

1. *Informed Consent* Before starting the user study the participants have to sign an informed consent. This form briefly explains what the study is about and clarifies that the participant is not obliged to finish the study. If the user terminates the study before finishing it, however, he cannot participate in the gift certificate raffle. Optionally, the user can agree with the anonymous publication of the:
   a) transcriptions of the study (recordings will be deleted after the study)
   b) filled out surveys
2. *General-Survey Before* At the beginning the participants have to fill out a general survey, where they have to judge their own knowledge on the topic of Internet security in general. For instance, they are asked whether it is easy for them to distinguish legitimate e-mails and websites from fake ones.
3. *Website-Survery Before* In this part of the user study the participants gets a list of screenshots of websites. The screenshots had been taken with the standard browser of an Android tablet. In total, the user is shown 16 screenshots, with 8 phishing and 8 valid URLs. The user has to decide whether he would enter confidential data on the shown website. Additionally, he has to encircle the part of the screenshot which was the primary reason for his decision. Then, the user has to indicate how sure he was about his answers on a Likert scale. Finally, the user is asked whether he knows the vendor of the website and whether he has a account there.
4. *Play App* After the "Website-Survey Before" the users get the smartphones in order to play the app. To save time, we skipped the introduction 2 part (access address bar) for the user study. The user has half an hour to play the app. After half an hour they are asked to put the smartphones aside. Then, we collect the smartphones and note the reached points in each level.
5. *Website-Survery After* After playing the app, the participants get a second website-survey. In this, all examples of the previous survey are included. Moreover, it contains 8 further website screenshots of which 4 have phishing and the remaining 4 have valid URLs.
6. *General-Survey After* Here, the participants are asked to fill out a form with questions to their demographics. This form does also contain questions related to the SUS and some other questions regarding their impression of the app.
7. *Certificates* At this point the official study is finished. We thank the users for their participation and award a "Golden Anti-Phish Certicate" to the best participant of the current group. All other users receieve a "Silver Anti-Phish Certificate". Next, the gift certificate is raffled.
8. *Debrief* Finally, an optional debrief follows, where the users can ask questions or provide their remarks in person.

10.2.1 Limitations

We have decided to conduct a study which compares the users' performance before and after playing our educational app. Our study design has some limitations we were not able to address for several reasons. This section discusses these limitations.

Behavior Change In our study the participants were not in their usual environment. Therefore, they likely behaved differently during our study. An alternative approach was to distribute the app to several participants and ask them to play it remotely. However, this has two major downsides: First, the user would have been remote and thus we would have less control. Second, testing the before and after app skills would have been difficult to realize.

Increased Attention At the beginning of the study the participants were told that the study dealt with phishing. Additionally, for the website-survey before and after they were explicitly asked to indicate whether the websites were phishing or not. That is to say, the user automatically increased his attention towards answering this question. In fact, this is not close to reality. Designing an in situ study, where the participant would have been in their usual environment and would

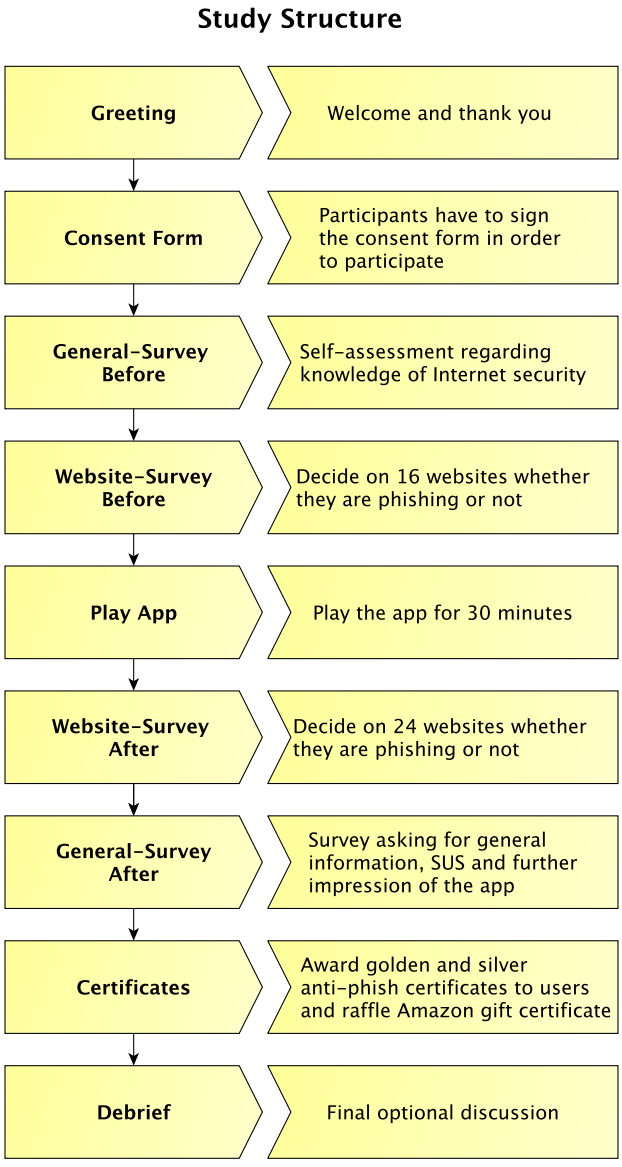


Figure 11: Our study structure

have not known about their participation, was not considered for time reasons and, most importantly, because such a design is ethically and legally questionable. For these reasons the post influence of the app could also not be tested.

Retention Our study design focuses on the present. For time reasons we did not have the possibility to repeat the "after app" scenario after some time has passed. Consequently, retention is an aspect which is not considered by our study.

Bias For the recruitment we endeavored to ensure that our participants are not close friends of ours. We rather seeked for friends of friends or even completely unknown persons by, for instance, hanging out our flyers and contacting several professors (cf. section 10.1). In order to further avoid biases the study should have been conducted by other persons than us. The fact that the participants were in contact with us, the app designers and developers, during the study might have caused a bias. Yet, we believe that the potential bias is negligible since the participants' feedback was not outstandingly good. Their feedback was generally positive, but some of the participants also criticized several aspects of the app (cf. section 10.5.4).

Despite the limitations of our study we believe that we could get a good insight into the effectiveness of our app. In the succeeding sections we will elaborate on the hypotheses we constructed in the previos section and analyze them.

10.3 Hypotheses

In order to evaluate the effectiveness and usability of our app we have formulated below mentioned hypotheses and measurements. As aforementioned, users are asked to encircle the area of the screenshot which was the primary reason for their decision. The coding of these markings, which are relevant for our measurement, can be consulted in section 10.4.

1. *Hypothesis 1 - Mistakes* After playing the app, the users make significantly less mistakes when deciding whether a website is a phish or not.
   Measurement: Correctly identified websites in "Website-Survery After" (phish or no phish) >> correctly identified websites in "Website-Survery Before"

2. *Hypothesis 2 - URL Based Decision* After playing the app, the users base their primary decision on whether a website is a phishing website or not significantly more often based on the URL compared to before playing the app.
   Measurement: Number of URL markings in "Website-Survery After" >> number of URL markings in "Website-Survery Before"

3. *Hypothesis 3 - URL Comprehension* After playing the app the user understands the importance of the domain of a URL as the only criteria to detect phishing websites
   Measurement: Number of marked URL domains in "Website-Survery After" >> number of marked URL domains in "Website-Survery Before"

4. *Hypothesis 4 - Good Usability* The app usability is above average.
   Measurement: According to [14] a System Usability Scale (SUS) > 68 can be considered above average usability.

10.4 Coding of Markings

In the previous section we stated our hypotheses and how we decided to measure them. For Hypothesis 3 it is interesting to specify what marking of what area we consider as separate option. Here, we provide an overview of our coding for possible markings in the website-surveys. These codings are relevant for the measurement of our hypothesis 3.

1. *None* Occasionally, participants did not mark or encircle anything of the screenshot. In our raw data this is coded as none.

2. *Favicon or Padlock* The marking of a favicon or padlock is trivially coded accordingly.

3. *Content* If anything else than the URL itself, a part of the URL, a favicon, or a padlock is marked then this is coded as content.

4. *Scheme* If the scheme or a part of the scheme in a URL is marked, this is coded as scheme.

5. *Host* Marking the host results in an according coding.

6. *Domain* In case a participant marks a domain or the substring of a domain this is coded as domain or domain substring accordingly. For the measurement of hypothesis 3, domain as well as domain substring markings are considered domains.

7. *URL* All other markings are coded as URL and measured as such.

10.5 Results and Analysis

This section presents our results and analyzes them. We start with discussing the representativeness of our participants and proceed with illustrating interpretation problems we faced while we assessed the website-surveys. Therafter, we analyze and present the results of our measurements to our hypotheses and proceed with some further exploration of our results. Finally, this chapter is closed with our discussion and conclusion.

10.5.1 Representativeness

In general we took care that the people we tried to recruit do not have extensive prior knowledge on this topic. For example, our flyer asked for non-specialists. Yet, we were not able to assure beforehand that we would only have non-specialist participants. In such a case we had to rule them out for our analysis afterwards. Unlike in our phishing survey (cf. section 6), where we ruled out electrical engineers or computer scientists in general, we did not generally exclude those from our final user study. The problem with the phishing survey was that it did not give us enough indication whether a particular participant was too skilled for our target group. Therefore, we had to imply that computer scientiests and electrical engineers are too skilled for our target group, even if this does not necessarily need to be the case in reality. In fact, there might be computer scientists or electrical engineers who can learn something from our app. In the contrary, in our final user study we were able to determine a user's prior knowledge more precisely with the aid of the website-survey before. Therefore, we did not primarily consider their course of study or field of work, but rather how well they performed in the website-survey before playing the app (cf. section 10.5.3). Finally, we made an effort to recruit participants who are not close friends of ours in order to minimize biases.

10.5.2 Marking Interpretation

As aforementioned, the study participants were asked to fill out a website-survery before and after playing our app. In this survey they had to indicate whether a given screenshot of a website is a phish or not. Furthermore, they were asked to encircle the area of the screenshot which was the primary reason for their decision. This section first shows some examples for the marking codes depicted in section 10.4. Afterwards, we illustrate you the challenges we had to face when interpreting the users' markings, as they were not always clear.

Marking Interpretation Examples As aforementioned, we categorized the possible markings of users into favicon, padlock, content, scheme, host, domain (substring), and URL in general. Figure 12 and Figure 13, for example, illustrate samples where the content or the padlock of a website was marked. Figure 14 and Figure 15 depict examples where the markings were interpreted as domain and domain substring. A sample where the host is marked is illustrated in Figure 16. And finally, Figure 17 shows an example of a marking (subdomain) which is coded as URL in general.

Interpretation Problem Examples While assessing and digitalizing our data we had to face some interpretation problems which we exemplify in the following. In such cases we had no choice but trying to interpret the samples as objectively as possible. Figure 18, for instance, shows a marking where the circle includes the content (youtube logo) as well as the

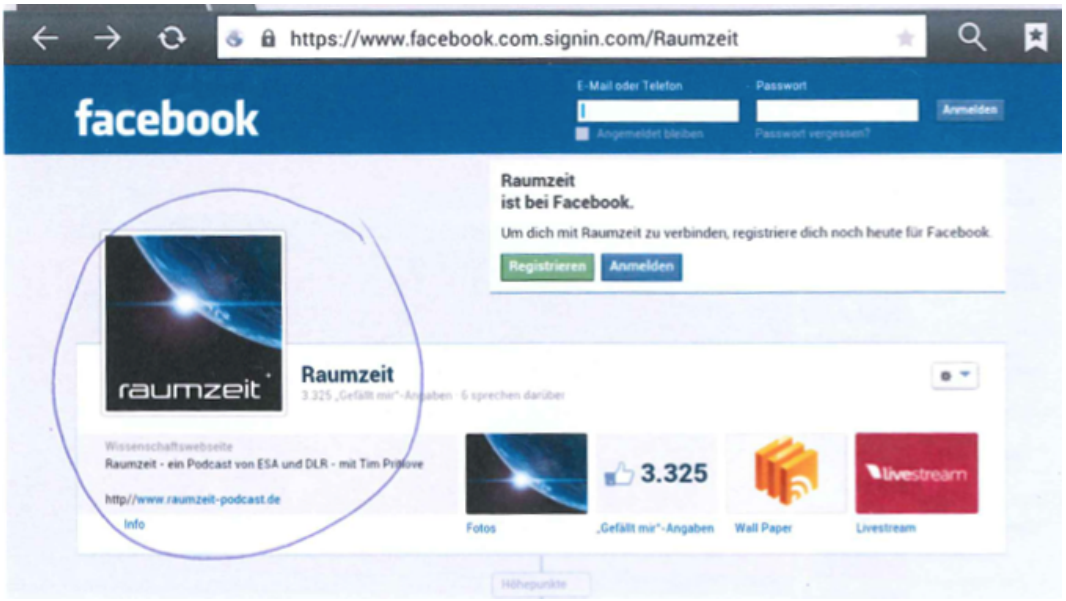


Figure 12: Content marked

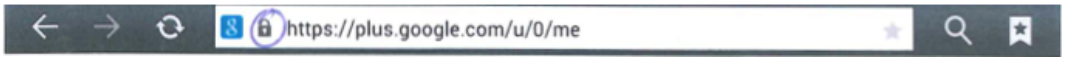


Figure 13: Padlock marked

host. For this sample, we decided to code this marking as host. The next examples in Figure 19 shows a marking where a subdomain and the domain is marked. When we faced examples like this we decided to code it as domain in case a subdomain is only partially marked, so that there is an indication that the user just did not make his markings clear enough. If the subdomain is obviously marked intendedly, i.e. clearly inside the circle, this kind of sample is coded as URL. In this case we see that the subdomain is inside the circle, for this reason we coded this sample as URL. A sample we came across often is a participant making two markings, even though we explicitly asked to mark only one area. In these cases we decided as follows: in case a marking is obviously striking due to a thicker circle, for instance, the more emphasized area is chosen for coding. Should both markings be kind of equal, we joint the markings and decided based on that. In Figure 20 a participant marked the scheme and the host separately. We cannot observe any emphasis on one of the markings. Therefore, we chose to code this sample as URL. Finally, there were samples where the user correctly identified a phishing website. However, instead of marking the domain as reason, what we expected from the participants, some users marked the attacked part instead. Figure 21 illustrates such an example. In fact, this reason is justified. Yet, we had to code such samples as URL since there is no clear way of defining the code for recognizing the attacked part of a URL.

10.5.3 Analysis of Our Hypotheses

In total 19 participants attended our study (one participant did not show up). As discussed in section 10.5.1 we did not rule out any participant beforehand. In fact, wer had to sort out two part two users for the results and analyses of our study:

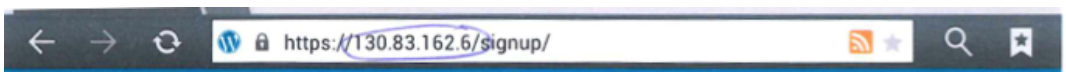


Figure 14: Domain marked

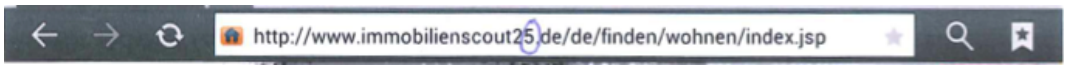


Figure 15: Domain substring marked

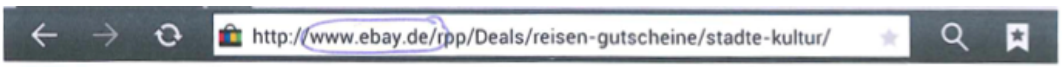


Figure 16: Host marked

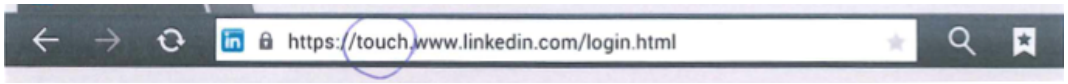


Figure 17: Subdomain marked, coded as URL

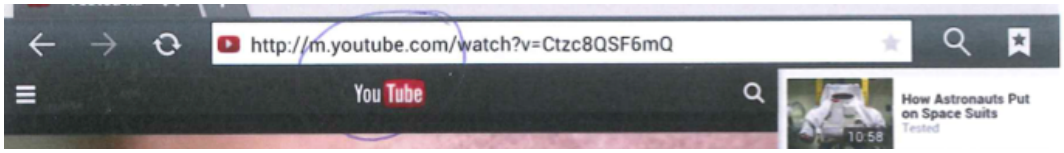


Figure 18: Host or content marked?

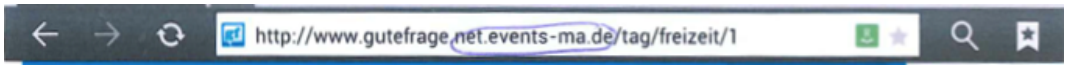
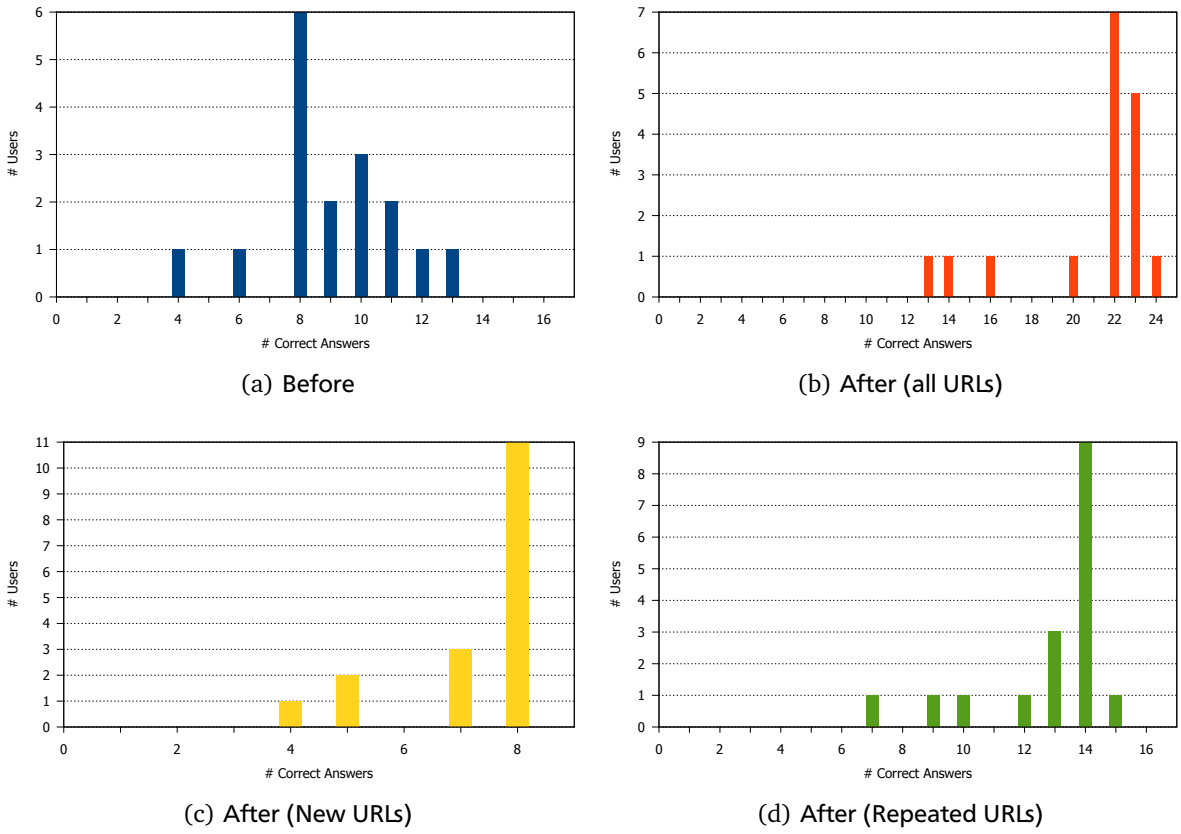


Figure 19: Domain or URL marked?

Figure 20: Scheme or host marked?

Figure 21: Attack marked



(a) Before

(b) After (all URLs)

(c) After (New URLs)

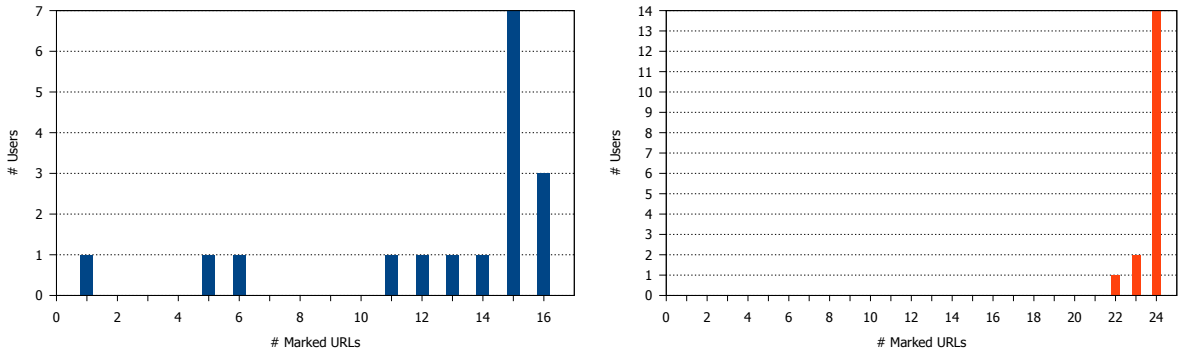
(d) After (Repeated URLs)

Figure 22: Correct Answers

1. *Outlier* **??** depicts the performance of our participants. It illustrates how many URLs were correctly identified by how many users. Evidently, there is an outlier among our participants. One user gave 15 correct answers to 16 questions. Thich means that he has too much prior knowledge on this topic and thus does not match our target audience. Therefore, this participant is not considered for our further elaborations.

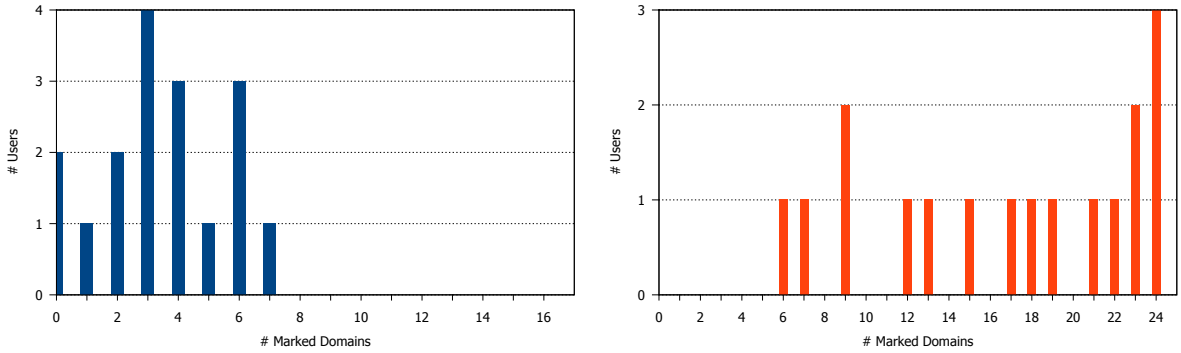
2. *Seriousness* Another participant obviously did not engage himself to play our app. During the 30 minutes of playing the app the user managed to complete the awareness part and level 1 (identify domain) only. More importantly, we saw the user playing around with the Samsung device instead. Since this user did not seem to take our study and app seriously we decided to sort him out for further considerations.

Hypothesis 1 Figure 22 shows the results of our study according to hypothesis 1. One can clearly see that the majority of the users identified more URLs correctly after using the app than before. While most participants correctly identified 8 out of 16, i.e. 50%, websites before they played the app, the majority gave correct answers to 22 out of 24 websites afterwards, i.e. 91.67%. One could argue that this increase is based on the fact that the examples are mainly the same in the website-survey after, i.e. the reason for their better performance is based on learning effects. Figure 22(c) however shows that the user also gave correct answers to most of the new URLs. Therefore, we assume the learning effects are negligible. In order to affirm our hypothesis we decided to apply the onesided Wilcoxon signed-rank test [88] with our 16 samples from the website-survey before and the same 16 samples from the website-survey after. Since we consider the learning effects negligible, we do not apply an alternative test against the 24 after URLs. Our null hypothesis is $H_0 : x_1 >= x_2$ and the alternative hypothesis $H_1 : x_1 < x_2$, where x_1 represents the number of URLs which were correctly answered before playing the app and x_2 represents the number of URLs which were correctly identified after playing the app. We computed the positive and negative rank-sums of $W_+ = 141.5$ and $W_- = 11.5$. The test statistic w is the minimum of W_+ and W_-, hence $w = 11.5$. As we chose $\alpha = 5\%$ as significance level and had 17 participants this results



(a) Before (b) After

Figure 23: URL marked



(a) Before (b) After

Figure 24: Domain marked

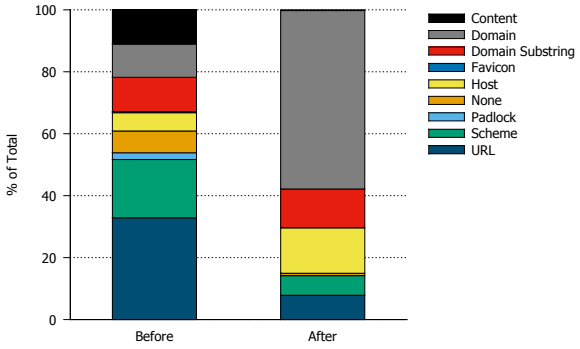


Figure 25: Marked parts of the screenshot before and after

in a critical value of 41. As our test value $w = 11.5 < 41$ the null hypothesis can be rejected and thus the alternative H_1 is accepted. Consequently, after playing the app the participants gave more correct answers than before. We are aware that we cannot fully rule out the possibility that there might be kind of learning effect. Yet, we are confident that these results cannot entirely be reduced to learning effects. Therefore, we believe that our app helped users to make improved decisions about the legitimacy of URLs.

Hypothesis 2 Figure 23 shows how many users marked the URL as their main source of decision. Most of the users already based most of their decisions on the URL before. Occasionally users marked the content or the padlock. However only 3 Users (17.65%) always marked the URL. Afterwards we see that most users (82.35%) always based the decision on the URL and only 3 Users made on or two mistakes. Therefore we think that our app emphasized their believe in basing their decision on the URL. We however think it is important to tell the user this to get uncertain or unknowing users to the same level as most of our participants. Also our app empathizes the importance of the URL by putting the main focus on it. We have decided against doing statistical test on this hypothesis. We believe it will most likely be rejected because the change from before to after is not very high.

Hypothesis 3 There is a general problem with one question in the websites-surveys. In the before survey we were not able to clearly ask the user to mark the domain when it was the base of his decision because we would have then primed them towards looking at the URL or even at the domain. This would have influenced the results of hypothesis 2. Since we could not formulate this question clearly, a user might have marked the whole URL even if his decision was based only on a small part (e.g. the domain) of the URL. Consequently, we were not able to clearly identify what the users' main source of decision was in the before survey. We were aware of this problem beforehand but saw no other option than formulating the question in such an open form. Afterwards, the user knew that they were expected to mark the domain. This can be interpreted as a change of question even if the literal question did not change. Therefore, we cannot apply any statistical tests on this hypothesis. Yet, we want to have a look at the results. None of the users marked the domain in most cases beforehand, in particular, 7 domains out of 16 URLs where marked at most by only one user, cf. Figure 24(a).

Figure 25 shows the distribution of the marked areas before and after playing our app. Obviously, afterwards most of the users marked the domain. However we are not able to compare that to the before values because of the changed question. Another interesting observation is that quite a number of participants marked the complete host (instead of the domain) in case of legitimate URLs, cf. Figure 26. One explanation for this might be that we did not ask the user to mark the domain of legitimate URLs except in level 1.

Hypothesis 4 According to the answers that the users gave in the After survey SUS section (Appendix D) we calculated a SUS of 83.1. This is above 68 which means that we can consider our app above average usable.

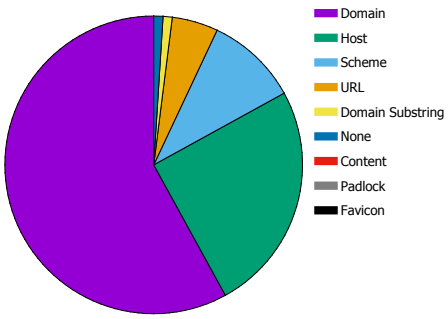


Figure 26: Marked URL parts of legitimate URLs

10.5.4 Further Exploration

Besides the results of our hypotheses the study yielded some further interesting outcomes. This section further explores these results.

Correct and Reasoned Answers Hypothesis 1 only refers to giving a correct answer to the question whether a website is a phish or not. Our results show that users did not only give more correct answers after playing the app. In addition to their correct answer they reasoned their answer appriopriately (i.e. marked the domain). In the website-survey before 37.5% of the URLs were answered and reasoned correctly. After playing the app 75% of the URLs were correctly identified and reasoned. Note, that reasoned means by our definition that the domain was marked in addition to giving a correct answer. However, a reasoning must not always rely on the domain itself. As discussed in section 10.5.3, for example, there were numerous participants who often marked the host of legitimate URLs. This is a correct reasoning for their decision, however by our definition it was not considered as such. Also, there were plenty of users who detected a phish and marked the attacked part instead (cf. Figure 21). By our definition this was also not accepted as correct reasoning. Hence, if we had expanded our definition of reasoning even more users would have correctly reasoned their decisions. However, there is the question whether expanding the accepted answers might also result in an increase of the correct answers and markings in the before survey.

User Opinions to App A part of our surveys tried to understand the users' opinions to our app. Section 3 of our survey-after (cf. Appendix D) contains the statements which the users had to assess with the aid of a Likert scale. Our worst score referred to whether the user was motivated by the spoofed e-mail to continue playing the app (average of 3.56) and whether the amount of texts was appropriate (3.53). We were beforehand aware that opinions of these two aspects might differ. This is also reflected by the individual answers. Many people strongly agreed with the statement that they were motivated and found the text amounts appropriate. However, there were obviously also people who disagreed with the statements. A reason for this might be, for instance, that they were already aware of the easiness of e-mail spoofing. All other statements were agreed with in average. The text legibility of our app texts received an average score of 4.7. Our study participants in average strongly agreed (4.82) that our app helped them to identify phishing websites in future. Yet, this requires a change in their behavior (checking the correct part of a URL) and retaining the obtained knowledge what we cannot check. Finally, the users intuitively understood our three lives scheme per level (4.411). All in all, our impression is that our app was well received by the participants.

Achieved Levels In average the participants reached level 7, which is higher than we had expected. We assume that some users started to roughly scan our texts for relevant information (i.e. looking at the examples) after they understood the importance of the domain. One user even played through the app in 30 minutes. In fact, this user has performed worse by the terms of our definition of correctness. The user had correctly identified 75% of the URLs before playing the app. Afterwards, he only had a score of 54.17%. The problem here was that level 9 deals with the difference of HTTP and HTTPS websites and that we did not expect the users to achive a level higher than 8. Therefore, we did not consider HTTP and HTTPS in our assessment. Users were asked to decide whether a website is a phish or not disregarding

the usage of HTTPS. The user who has achieved level 9 however was explained the difference of HTTP and HTTPS, additionally, he was asked to reject HTTP sites in general for this game. Hence, this user responded to the website-survey after respectively: the participant generally rejected HTTP sites whether they were phishing or not and thus the user performed worse afterwards by the terms of our definition of correct.

HTTPS and Padlock When assessing the before website-surveys we had the impression that many participants where aware that they should look for either HTTPS or the padlock. Yet, they seemed not to be aware of the fact that the use of HTTPS does not necessarily mean that the website is trustworthy in general. In fact, there may be phishing websites using HTTPS. These participants fell for such websites in our survey and are likely to fail to such attacks in reality. Therefore, we think it is important to move the level which is dealing with HTTPS to an earlier level since we cannot be sure whether an app user of ours plays until level 9. This aspect should definitely be covered in future work in our opinion.

During playing the app, the participants had a slip of paper for notes they wanted to make considering the app. In the following we outline the main results of these slips of paper. Note that we did not ask the users to write down something specific. We merely asked them to write down what they thought, i.e. there might be more participants who agreed with some of these points below but just have not explicitly written it down. In the following we consider the notes and suggestions of all participants.

Scrolling of URL In addition to deciding whether a URL is a phish or not, the user has to face two more challenges. First, the font size gets increasingly smaller in higher levels, until it eventually is approximately the same size of the Android standard browser. Second, the URL is displayed in a horizontal scrollbar so that he has to scroll the URL to the right in order to view the beginning of it, just like it is the case in browsers. 4 of our 19 participants (21%) found this disturbing and said it hindered them from analyzing the URL reasonably. Well, this is exactly what we wanted, since the behavior in the browser is the same, and users should practice it. We assume that the users would not have noted this in this extent in case they would have had to complete introduction 2 (access the address bar), because then they would have understood why we do the URL scrolling. Yet, we think after a couple of levels the users should have understood that they have to scroll the URL, in the game as well as in the browser, so one might consider to eliminate after some time.

Unknown Services We have mentioned the problem of unknown services in section 7.4.2. As we were afraid there are in fact services which are not familiar to several users. 4 out of 19 participants (21%) mentioned this problem. Even if we tried to make use of the most popular services with the aid of Alexa's [20] ranking we cannot assure that all used URLs are known by all users. One idea to approach this challenge might be to provide an question mark button in addition to the check mark (no phish) and cross mark (phish) buttons. When a player clicks on this button, he can be told whether the given URL is a phish or not and why. From this action the user would neither profit nor would he lose any points or lives. Yet, we do not think that this is a major issue, since the users got at least until level 4 and most of them achieved even higher levels. We are confident that the app in its current state is already implicitly able to teach the users about the legitimacy of unknown services. After facing unfamiliar URLs and making or not making mistakes they will eventually learn whether to trust a service or not.

Question to Data Entry Originally, the website-surveys before and after asked the participants whether they thought a given website screenshot was a phishing website or not. After our test iteration of our study, however, we decided to change the question towards asking the user whether they would enter their personal data into this website in order to create some context for the user. As we told the user that this study was particularly about phishing and that their task was to detect phishing websites in the website-surveys before and after, we believe that it was clear what we were asking for with this question. Yet, 3 participants (15.8%) noted that the formulation of this question was unclear, i.e. there might have been participants who selected "no" even if they did not think it was a phishing website, but they would generally not enter their data in this website.

Explanations and Comprehensibility 4 participants (21%) stated on their slips of paper that they found the explanations of the app very good and easy to understand. 1 of these 4 participants, however, added that there is partially much text to read. Another participant (not under those 4) noted that there is too much and long text in general.

App Structure We did our best to develop an app which is consistent and well-structured. 3 of our participants (15.8%) confirmed our intentions by stating that they found our app well and clearly-structured.

Button Positioning The positioning of our app buttons during the game are as follows: The left bottom corner has a check mark which represents that the user thinks the displayed URL is not a phish. The right bottom corner has a cross mark which means that the user thinks the displayed URL is a phish. After clicking on either of these buttons in the write bottom corner another button appears (where the cross mark usually is) which either is the continue or the verify button (depending on whether the user has to select the Who-Section or not). 2 of our participants (10.5%) indicated that the positioning of the buttons in the right bottom corner are suboptimal. The problem here is that accidentally double clicking, for example, the continue button in the right corner results in rejecting the next URL even if the user might not have intended to. Even if only 2 participants explicitly criticized this aspect we believe that this is a legitimate point. In fact, the positioning of the two buttons continue and verify should be different from the one of the cross mark. This is an aspect which should be targeted in future work.

Repetition Repetition is an important element of our app. In every level introduction we briefly repeat the so far learnt parts of a URL (with a graphic) and the different attacks the user has seen until this point. We also make use of repetitions during the exercise rounds, every level contains at least one exercise from the previous level. 2 of our participants (10.5%) explicitly indicated that our repetitions made them feel more confident and safer.

External Links In the main menu of our app we have a button “More About Phishing” which leads to a list of external links to various websites about phishing. As it is often easier to find websites on a specific topic in English, some of these websites are in English as well. 2 participants (10.5%) indicated that they did not like it to be led to an English website and would have expected to be forwarded on a German one. This aspect is also a point which is worth to consider for future work, since we cannot expect our audience to have knowledge of the English language. An idea to approach this might be to provide in-app additional info. That is, instead of linking to external websites, the app itself could provide additional categorized information in German. This would solve the problem with the language and at the same time the additional information would fit to our app layout and design.

Amount of Examples Our app starts with a small sample of URLs users have to decide on. In every level the sample size increases as the number of possible attacks increase. 2 participants (10.5%) found that our sample size was to big. This might have been the result of the fact that the users had to play the game for half an hour at a time. Yet, re-thinking the number of URLs per level might be reasonable, but one has to consider that there have to be enough phishing URLs for the repetition as well as the new attack in each level.

Mail Not Received The first part of our app includes sending a spoofed e-mail to oneself in order to increase the awareness of how easy faking e-mails is, even for unexperienced users. 1 participant did not receive the e-mail he tried to send to himself. He skipped this part by clicking on the according button we had added for such cases. All other participants received the spoofed e-mail.

Further Suggestions In their notes some users made several suggestions which we found interesting and would like to mention here. 2 participants stated that the attacks in level 2 are not very challenging and very obvious. Phishing URLs are of the following kind in this level: “http://www.ebay.de.phisher.com/login”. We had done this intentionally, however maybe the app could in fact start with for example nonsense attacks or just reduce the number of samples in this level.

Also some users suggested us to visually distinct new learning content from repetition so overflying the information part becomes easier. We had already tried that by displaying a warning sign every time a new attack was introduced. However, after this sign there still was some repetition occasionally. Therefore, the emphasis of new and relevant learning content in the lesson parts should be improved in future work.

Another participant noted that the repeating sentences “Very good. You are now a bit safer”, for example, when the user correctly identified a phish, were kind of repetitive and not motivating enough. This is ultimately related to our issue with the Law of Effect in section 8.6.1. There we stated that the texts and icons of the app should vary according to

the user's performance and the degree of difficulty in order to achieve higher and better positive feelings. This is an additional point which is worth to consider for future work.

A participant would have wished more "feedback" on his progress in general. He missed statistics, highscores or other forms of long-term feedback. As we mentioned in section 8.6.1 and section 8.6.2 feedback is essential for learning as well as games. For this reason, this aspect is something which should be enhanced in future.

A further interesting point is that one of our participants noted that he thought our app was appropriate for schools. From this we infer that the participant found our app easy to understand and that he thought pupils were capable of understanding and playing our app. In fact, this is a very good and important point. People might not want to learn about phishing on their own, especially pupils would probably not think about that. However, if computer security is a part of the education plan and if pupils have to learn something about, for example, phishing by playing the app, a far larger audience can be reached and educated on this topic.

Finally, one user indicated that the design of our start menu could be more appealing. This is a justified criticism. In fact, we did not put our focus in design aspects. To make the app's outer appearance more appealing design aspects could be considered for future work.

10.5.5 Legibility Index

Comprehensibility is of major importance for our app. Each level starts with a lesson on a certain way phishers make use of in order to delude people. These lessons aim to educate the user by describing those phishing attempts. Hence, it is of utter importance that our explanations are both thorough as well as comprehensible. After we had conducted user tests and included their corresponding valuable feedback (cf.section 9.2) and also received good feedback from the final study (cf. section 10.5.4), we now want to assess the comprehensibility of our texts with the aid of a statistical method. Several approaches to assess the readability of a given text exist [38, 30]. These approaches consider, among other details, the average word and sentence length and therefore are language dependent and usually are designed for the English language. However, the Flesch-Reading-Ease [30] is a legibility metric that outputs a numeric indication for the readability of the input text and was also adjusted to operate with the German language by Toni Amstad [**?**]. The readability of a German text is computed as follows, where ASL represents the average sentence length and ASW the average number of syllables per word:

$$\mathrm{FRE}_{\mathrm{German}} = 180 - \mathrm{ASL} - (58.5 \times \mathrm{ASW})$$

Several auxiliary tools exist that take a regular text as their input and return a legibility index as their output [49, 83, 75]. All of these tools delivered different scores because they obviously had varying algorithms to determine syllables, sentences and even word boundaries. Therefore, we approached as follows: First, we took the explanation part of each level as the input for all of the three tools. Second, we extracted the amount of sentences, words, and syllables from the returned values, as depicted in Table 2. We did not rely on the returned legibility indexes, because the tools seemed to use slightly different formulas. Third, with the extracted information we were able to compute the German FRE index ourselves in a consistent way through all of the three tools. Ultimately, we derived a final index value of 62 by averaging the resulting index values of each tool. Given a scale from 0 to 100, where an index of up to 30 indicates an academic level and 90 and above is considered easy to understand, an index of 62 is considered as reasonably comprehensive for teenagers [4]. Regarding our target group, this is a good result and confirms the comprehensibility of our explanations also statistically.

10.6 Discussion

Our hypotheses stated that our app can increase three major skills of the user. First, we wanted to give the user the ability to detect phishing URLs. This goal can be considered achieved as we could show that the increase is significant. Even though there might be some learning effect we are confident that this increase is not mainly attributable to that. The second and third hypotheses focused on the reasoning behind the user's decision. We wanted to show that the user is

	Leichtlesbar.ch	Stilversprechend.de	Fleschindex.de
# Sentences	426	604	400
# Words	4616	4760	4762
# Syllables	8305	8658	9235
Legibility Index	64	66	55

Table 2: Sentences, words and syllables of our texts outputted by different tools [49, 83, 75]

aware of the fact that the content is no source of evidence against a phishing attack. Our results suggest that most of the users already feel that the URL is important beforehand but some additionally consider the content. Thus, the change in user response was not high enough to measure our hypothesis 2. The third hypothesis said that the users understand the structure of URLs better after playing our app. As we discussed above, we had concerns about the design of the question that was intended to test this hypothesis. The question has to be regarded as changed after playing the app. Due to this, we could not consider the before-markings and could not argue on any findings about this hypothesis.

In addition to testing the hypotheses we got overall positive feedback from the users. Most of them had the feeling they learned something from the app. Some participants even contacted us asking about the release date of the app because they wanted to give it to their relatives.

Yet, there are still some improvements that should be considered when someone develops a next version of the app. First of all, we think that the part talking about HTTP and HTTPS should be moved to an earlier level. We saw many users that marked the scheme or the padlock in the URLs in the website-survey before. This seems to be more important for the user than the URL itself because we had several users who fell for a phish and marked the scheme or padlock as reasoning. As we cannot assume that users finish our app within one day, or even finish it at all, we think that it is important to clear that misunderstanding earlier. The second improvement that a following developer should include is to restructure the app texts in such a way that the user can clearly identify repeating and new parts. We think this aspect is not as important in real life compared to study situations because the app is not designed to be played a long time in a row. This would increase the importance of the repetitions and make such a separation less important but this can be improved. Last, it might be a good improvement to modify the app behavior depending on the user skill and performance. This might prevent the user from getting bored. This includes, for instance, vary feedback texts and icons depending on the user's performance as well as the degree of difficulty. Another example is that the app could skip the proof part (identify Who-Section after detection of a phish) in case there is an indication that the user understood how to do it. A simple form of that is already implemented. The proof is shown up to a predefined level. Yet, we think it is more reasonable to base the skipping and other possible extensions on the user performance.

To conclude our findings despite some possibilities for improvement we can say that the app helped most of the users to protect them from phishing. This means we overall achieved the goal of the app.

11 Conclusion, Recommendations and Future Work

This chapter provides a short summary of what we achieved in the scope of this thesis and some further concluding remarks. We also present a short list of recommendations regarding the design of security education games based on the lessons we have learned. Finally, we have a short outlook on future work.

11.1 Conclusion

The objectives of this thesis were twofold. First, we aimed at increasing users' security awareness so this may hopefully result in a change in their security-related behavior. Second, we focused on the education of users with regard to the detection of phishing URLs. Our app is supposed train the user to achieve the required capability of correctly parsing URLs and thus identifying phishing websites. This capability will hopefully help him to defend himself against phishing attacks.

To achieve these goals we developed an anti-phishing education quiz based game. Our app targets the awareness increase by actively let the user spoof an e-mail and exemplifying him that a link does not necessarily lead to the target that it displays. By letting the users practically experience this, we hope to increase the intensity of their learning experience (cf. Principle of Intensity in section 8.6.1) resulting in a better and higher learning performance and motivation. We are aware that this part of the app might not be interesting and motivating for some users. This is especially reflected by the answers to our question in the survey-after, where we ask whether this part motivated them to continue playing the app. The answers to this question are rather dispersed and thus result in an average value of 3.611 out of maximal 5 which would indicate that the user absolutely agrees. Yet we think it is an important part of the app since there still seem to be users who are not aware of the facts that are taught in this part (the median is 4 out of 5).

After motivating the user with practical examples and increasing his security awareness the actual game starts. The game itself consists of levels with introductory parts followed by practical exercises the user has to solve in order to show he has understood the learning content. Initially, in introduction 2 (access address bar and view complete URL) and level 1 (URL basics and domain identificationzich), basic knowledge is covered which is required for the succeeding levels and especially for the detection of phishing URLs on the smartphone in general. In particular, introduction 2 might be not challenging enough for some users as they might be already well-skilled regarding smartphone functionalities. Yet, this introduction and exercise is indispensable as there might be users who do not know this.

In levels 2-9 the user is introduced to various attacks a phisher might apply. These attacks get increasingly sophisticated and harder to detect with higher levels. Finally, in level 10 the user gets further final remarks, such as the discussion about legitimate URLs which might appear fraudulent but in fact are not, or some input to extended validation certificates and a link for further information. We generally aimed at using introductory texts that are simple and easy to comprehend. We think we achieved this goal as our participants agreed to this (cf. section 10.5.4) and the legibility index further confirms it (cf. section 10.5.5).

The study outcomes suggest that there is a possitive effect which is likely resulting from our app. Our participants clearly gave more correct answers (is a given website a phish or not) after playing the app compared to before. Before playing the app most users identified 50% of the 16 websites correctly. After playing the app the majority of the participants gave correct answers to 91.67% of the 24 websites. We assume the learning effects are negligible as the distribution of the correctly answered new URLs is almost identical to the old one (cf. Figure 22(c) and Figure 22(d)).

The results of our second hypothesis exposed that the users already knew it was important to look at the URL before playing the app. Yet, they obviously did not know where exactly too look at as their result in correct identifcations show. Even if there were many participants who looked at the URL already before playing the app there is a significant difference in the following aspect: Before playing the app only 3 (17.65%) users always marked the URL. In contrast to that, after playing the app most, i.e. 14 of 16 users (82.35%), always marked the URL. Evidently, our app was able to emphasize their believe in basing their decision on the URL rather than the content or anything else.

The measurement of our third hypothesis was questionable. The problem is the following: after playing the app we virtually changed the question when asking the user to mark the area of their primary reason for their decision. With the app we primed them to mark the domain clearly. Before using the app the users might have meant the domain, but

just did not clearly mark it because they did not know what exactly they were expected to mark. Still, we analyzed our results on this question and found out that more people base their decision on the domain after playing the app.

All in all, we can say that our app has a positive effect on users and that it helps them to better identify phishing URLs. However, the question to ask here is will this app actually help users change their behavior in the Internet and make them look at the URL even if it is only occasionally. This is an aspect, which we were not able to address within the scope of this work Finally, the study conducted cannot show how the users, for example, retain the lessons learned from our app. Such considerations remain open for future work.

11.2 Recommendations for Security Education Games

Technical solutions are not 100% accurate at detecting phishing attacks. The education and training of users offers a complementary approach to these systems. Based on our gained experience, we present design principles that we recommend to consider when designing a security education game:

Principles of Learning Since education games do not primarily aim at entertaining the user, but simultanously at educating him, it is important to take the priciniples of learning into account. These principles state under which conditions learning performance is increased (cf. section 8.6.1). We consider it especially important to rely on the principle of exercise, which states that training, repetition and feedback is crucial for good learning performance.

Game Techniques If the education is supposed to follow with the aid of a game it is relevant to regard essential game techniques (cf. section 8.6.2). In fact, game techniques are closely connected to learning principles. They provide in-depth elaboration on how basic learning principles are achieved with games.

Simple and Short Text Education implies that some sort of text is present in some way. The users to be educated may come from different fields. While some users might be more skilled and might be able to handle complex texts on security-related topics others would probably get discouraged by such. Skilled users are likely capable of acquiring knowledge on security topics without problems. Security education should mainly address those users who are overwhelmed by such texts and information. For these users it is important to provide simple texts which are easy to follow and not too long. The longer texts are the more likely it is that the users will skip text parts which might be important or even stop reading it.

Precise Phrasing The texts should be formulated precisely. If a text is not precise this might lead to misinterpretations and thus to mistakes. One should take care that there is as less room left for misinterpretation as possible.

General Validity There is a range of potential learning content which might be important to communicate to the users. Yet, there is the problem of general validity. It is important to consider that, for example, aspects that apply to system A, for example an Android device, do not apply to system B, for example another version of the same Android device. Therefore, in some cases it might not be easy to transfer knowledge about A to B (for example the browser functionality). For this reason it is relevant to consider whether one wants to educate the users about aspects which are generally valid among several systems or whether the education focuses on specific systems which ultimately would result in restricting the target audience to those who use that specific system.

11.3 Future Work

This section deals with a prospect on future work for our Anti-Phishing Education App. In particular, we present ideas that might be beneficial and which we were not able to address and realize due to time and resource limitations.

Comparative Study A comparative study might be interesting, as the authors of Anti-Phishing Phil [79] did it. They conducted a study with three different conditions, a group that consulted general tutorials from the Internet, a group who learnt from tutorials based on Anti-Phishing Phil and another group who played Anti-Phishing Phil itself. An interesting comparison to consider might be our app with Anti-Phishing Phil or any other anti-phishing app and general tutorials.

Study on Retention For time reasons knowledge retention is an aspect we could not address in our study. Yet, we think it is important to consider this in future work. The question to ask here is how well users retain the knowledge they obtain from our app compared to other sources, for instance.

Embedded Training In section 3 we argued that exploiting the teachable moment might result in good motivation for a user to do something about his lack of knowledge regarding phishing and possibly result in better retention. Yet, a drawback of embedded training is that its landing pages are not able to provide detailed and extensive information on the topic since users would be discouraged and leave the page. Providing detailed and extensive information can be addressed by a game playfully. Therefore, an idea is to combine embedded learning with an educational game. Here the user would be forwarded to the landing page with the most important information in case he falls for a simulated phishing attack. On this page the user can be provided with the most important information and a link to an educational app, for example ours, where he can optionally get more detailed information. With the game the knowledge can be obtained step by step by playing the game. However, there remains the problem of raised legal issues.

Malicious Downloads With our app we did not target the possibility of downloading malicious software when a user clicks on such a link. However, we think this is an aspect which should also be targeted in future work since malicious software can also cause harm. For example, the user could be told at some point that he should never open downloaded files he did not intend to download. Instead he should immediately delete them.

Certificate Validation We do not address the validation of certificates with our app. We also do not tell the user, for example, that he should at least not do banking in case a certificate is broken. Such general suggestions might also be considered for future work.

Data Economy Another relevant aspect we did not cover in our app is data economy. We tell the user to type in their data only into websites they are sure about their legitimacy. However, we do not tell the user to think about the specific data they are asked to enter. Users should re-think whether the required information is actually needed. This should also be trained by for example asking the user “A lottery site is asking for your data. Which data would you provide?”. As this approach would require a complete different UI this type of learning content remains for future work.

Consequences Initially, our idea was to display the consequences for falling to a specific phishing URL (matching a certain website category). For time reasons, we had to delay this for future work. We still think that this kind of information is relevant for the user as it illustrates him on which websites he should especially take care, for example, on banking websites.

Increase Immersion Immersion is an important game element (cf. section 8.6.2). We believe our app has space for more immersion. For example, an appealing story around our quiz game could be added in order to mesmerize the user.

Increase Effect In section 8.6.1 we have introduced the Principle of Effect, specifically, the law of positive feelings. We believe the user’s positive feelings can be increased by providing more variable feedback, instead of saying the same sentence for the same outcomes. For example, the praises and compliments when a user does well could vary depending on the degree of difficulty. Also, the final screens for finishing a level can vary.

Top-Level Domain Attacks It might be reasonable to explicitly tell the user, that he should not only look whether the second-level domain is exactly as he expects but also the top-level domain. The app only implicitly states to look at the top- and second-level domain together. We recognized that users might misinterpret this (problem of precision). Therefore, this addition should be realized for future versions.

Performance Dependent App Behavior Currently the app is quite static. More dynamic behavior could be added in future versions. For example, the part where the user has to show the domain in case he found a phish might get tedious after some time. To approach this problem the app could stop asking for the domain after the user identified the domain correctly 10 times in a row, for example. After this point the app could occasionally ask the user to identify the domain and depending on his performance re-introduce it.

A E-Mail template

```
<html>
  <head>
    <title>Anti Phishing Education</title>
  </head>
  <body>
    <p>Dies ist eine automatisch generierte E-Mail im Rahmen einer Anti-Phishing
       Education App. Falls diese nicht angefordert wurde, bitte ignorieren.</p>
    <p>Ansonsten geht es hier weiter:</p>
    <p>Wie du im Absender siehst, hast du dir gerade selbst eine E-Mail mit gefä
       lschtem Absender geschickt. Hier ist außerdem dein Freitext:</p>
    <p>{$usermessage}</p>
    <p>Für einen Angreifer ist es ebenso einfach automatisierte E-Mails mit gefä
       lschtem Absender und Inhalt zu verschicken. Meist enthalten diese einen Link zu
        einer Webseite, genau wie diese E-Mail.</p>
    <p>Um mit der App fortzufahren, klicke auf den folgenden Link.</p>
    <p><a href="http://pages.no-phish.de/maillink.php">http://www.google.com</a></p>
    <p>Viele Grüße,</p>
    <p>Dein NoPhish Team</p>
  </body>
</html>
```

B URL Generation

While playing the app the user is presented with URLs that he has to categorize as phish or valid. While reviewing the previous works and games in this area we found that many of them use a fixed set of examples. On some games this set is very small and therefore you are always confronted with the same URLs. As we layed out in section 7.3 we want to teach the user how to detect phishing URLs in general. To accomplish this goal we think that it is essential that the user sees as much different URLs as possible so he can build his own mental model. Therefore we decided on generating URLs rather than composing a fixed list. We will lay out the general process here and cover interesting parts of it in the following sections.

B.1 Example URLs

To present attacked URLs to the user we found int most realistic to take valid URLs and apply attacks on them. Therefore we needed a set of valid URLs. To build this set we used Alexa [20] to find the top 100 domains for german users. We then went to each of these sites and by navigating tried to find 6 URLs for each domain. We tried to find some short and some long URLs.

Generate Attacks for Level When starting a new level we generate a list of Attacks that we want to show the user.

Select Valid URL When we want to show a new URL to the user we first randomly select a valid URL from the before mentioned set.

Apply Generator Then we apply a generator to the URL that does not invalidate the URL but modifies it.

Apply Attack After that we select a random attack from the previously build list and apply it to the URL.

Repeat In some situations we need to try again.

B.2 Generate Attacks for Level

The types of URLs the user is presented is dependent on the level. Each level introduces on or more attacks. Which attack is introduced in which level is layed out in section 8.5. In general the URLs of each level n are distributed as follows:

Total number of URLs	u	$6+2*n$	starting with 6 URLs each level has 2 more URLS.
Number of Phishes	p	$u/2$	Half of the URLs are phishes.
Number of repeats	r	$\lfloor p/2 \rfloor$	Half of the phishes are repeats.

Table 3: distribution of URLs per level.

The repeats are always one attack from each previous level. The rest of the repeats is filled up randomly.
There are two main exception to these rules:

Level 1 In Level 1 the game is modified in the form that the user is only presented with valid URLs and has to select the domain. To prevent boring the user in this level we only present 5 URLs. None of them is a phish.

Level 1+2 The first level that contain repeats is level 3 because level 2 is the first real game level.

The generated list of attacks also contains a special attack that does no real attack. This is to simplify the URL generation. When we generated the list of attacks we save it for later reference.

B.3 Apply Generator

We were unsure if we still have enough valid URLs so we prepared a way to automatically modify the URLs in such a way that they could still be valid URLs. Some Ideas where to add subdomains or path strings to the URL. Query or fragments are also possible. We later found out that it is currently not needed to implement generators because there are a lot of URLs in our set. If however we will some time in the future find out that these URLs are not enough we have this scheme in place.

B.4 Apply Attack

After we generated a valid URL we chose a random attack from the previously build set of attacks and apply it to the URL. With this we also store which attack we currently applied. This is important when the user is failing this round. In this situation we will simply readd this attack to the set of attacks.

B.5 Repeat

There are combinations of base-URL and attack where the attack doe's not alter the URL. Therefore it would be impossible for the user to detect the Attack and he will be confused and might stop using the app. In this situation we repeat the whole URL generation process until we find a matching URL.

C Presurvey Form

Fragebogen zu Phishing im Internet

Im Rahmen unserer Studie möchten wir das Bewusstsein über Gefahren im Internet erforschen. Wir würden uns freuen, wenn Sie sich die Zeit nehmen könnten, uns ein paar Fragen zu beantworten. Die Daten werden natürlich anonymisiert gespeichert, sodass kein Rückschluss auf Ihre Person erfolgen kann.

1. Generelles

Zuerst ein paar generelle Fragen zu Ihrer Person.

1.1. Ihr Alter.....	_____			
1.2. Ihr Geschlecht	☐Männlich	☐Weiblich	☐Anderes	
1.3. Beruflicher Abschluss				
☐Lehre	☐Meisterprüfung	☐Hochschulabschluss	☐Keiner	
1.4. In welchem Bereich arbeiten/studieren Sie zurzeit?	_____			

2. Internetnutzung

Es folgen einige Fragen zu Ihren Erfahrungen im Internet.

2.1. Wie häufig sind Sie online?						
☐Ständig	☐Mehrmals täglich	☐Täglich	☐Mehrmals wöchentlich	☐Seltener		
2.2. Welche Anwendungen des Internets verwenden Sie?						
☐E-Mail	☐Browser	☐Banking	☐Medien (Audio, Video)	☐Spiele	☐Soz. Netze	☐_____
2.3. Verwenden Sie ein Smartphone?						
☐Nein	☐Ja, Android	☐Ja, Apple	☐Ja, _____			
2.4. Wenn ja, welche Anwendungen verwenden Sie auf Ihrem Smartphone?						
☐E-Mail	☐Browser	☐Banking	☐Medien (Audio, Video)	☐Spiele	☐Soz. Netze	☐_____
2.5. Wie viele Werbemails bekommen Sie pro Woche ca.?						
☐< 10	☐10-20	☐20-50	☐50-100	☐> 100		
2.6. Wie viele E-Mails, die Sie nach persönlichen Daten fragen, bekommen Sie pro Woche ca.?						
☐< 10	☐10-20	☐20-50	☐50-100	☐> 100		
2.7. Ich informiere mich über Gefahren im Internet und wie man sich davor schützen kann.						
☐Nein	☐Ja, selten	☐Ja, gelegentlich	☐Ja, regelmäßig			

3. Einschätzungen

Im Folgenden sollen Sie entscheiden wie sehr Sie den entsprechenden Aussagen zustimmen.

Bitte bewerten Sie zwischen 1: „Trifft voll zu“ und 5: „Trifft überhaupt nicht zu“	1 trifft zu _ 5 trifft nicht zu
3.1. Ich kenne mich gut genug aus, um den Gefahren im Internet aus dem Weg zu gehen.	1 □ 2 □ 3 □ 4 □ 5 □
3.2. Es fällt mir leicht legitime von gefälschten E-Mails zu unterscheiden.	1 □ 2 □ 3 □ 4 □ 5 □
3.3. Sicherheit im Internet bezieht sich ausschließlich auf Finanzanwendungen.	1 □ 2 □ 3 □ 4 □ 5 □
3.4. Meine Daten gebe ich im Internet nur an Anbieter weiter, die ich gut kenne.	1 □ 2 □ 3 □ 4 □ 5 □
3.5. Mir persönlich ist es egal, was mit meinen Daten im Internet geschieht.	1 □ 2 □ 3 □ 4 □ 5 □
3.6. Für den Schutz vor Gefahren im Internet ist jeder selbst verantwortlich.	1 □ 2 □ 3 □ 4 □ 5 □
3.7. Ich vertraue E-Mails, die von mir bekannten Personen kommen.	1 □ 2 □ 3 □ 4 □ 5 □

4. Phishing

Die folgenden Fragen beziehen sich auf einen Angriff, der allgemein als „Phishing“ bezeichnet wird.
Dabei wird versucht dem Nutzer Daten durch gefälschte E-Mails oder Webseiten zu entlocken.

4.1. Folgende Dienste sind durch Phishing gefährdet.

☐E-Mail	☐Browser	☐Banking	☐Medien (Audio, Video)	☐Spiele	☐Soz. Netze	☐_____

4.2. Folgende Daten sind durch Phishing gefährdet.

☐PIN/TAN	☐ Kreditkartendaten	☐ Zugangsdaten	☐Pers. Daten (z.B. Geburtsdatum, Adresse)

5. Anti-Phishing Anwendung

Ein Teil der weiteren Studie soll die Entwicklung einer Smartphone Anwendung sein. Diese hat das Ziel den Benutzer über die Gefahren von und Abwehrmechanismen gegen Phishing-Angriffe zu informieren. Zur Vorbereitung dieser Entwicklung interessieren wir uns für Ihre Vorstellungen zu einer solchen App.

Bitte bewerten Sie zwischen 1: „Trifft voll zu“ und 5: „Trifft überhaupt nicht zu“	1 trifft zu _ 5 trifft nicht zu
5.1. Passend zum Begriff Phishing fände ich ein Spiel mit einem Fisch lustig.	1 □ 2 □ 3 □ 4 □ 5 □
5.2. Um Wissen zu vermitteln ist ein Textbasiertes Programm am sinnvollsten.	1 □ 2 □ 3 □ 4 □ 5 □
5.3. Mit Frage-Antwort Spielen macht mir Lernen am meisten Spaß.	1 □ 2 □ 3 □ 4 □ 5 □
5.4. Zum Auflockern von Wissensvermittlung sind Comics ein gutes Mittel.	1 □ 2 □ 3 □ 4 □ 5 □
5.5. Zum Verfestigen des Gelernten sind Übungen unerlässlich.	1 □ 2 □ 3 □ 4 □ 5 □
5.6. Mich begeistern häufig Lernspiele.	1 □ 2 □ 3 □ 4 □ 5 □
5.7. Wichtig ist, dass mich ein Lernprogramm immer wieder überrascht.	1 □ 2 □ 3 □ 4 □ 5 □
5.8. Bei einer Anti-Phishing-App müsste ich für mich sofort feststellen können, dass es mir etwas bringt.	1 □ 2 □ 3 □ 4 □ 5 □
5.9. Gut fände ich eine Trennung zwischen dem Lern-Modus, dem Übungs-Modus und dem Test-Modus.	1 □ 2 □ 3 □ 4 □ 5 □

5.10. Folgende Idee/Anmerkung/Vorstellung habe ich für Ihre App.

5.11. Was müsste das Spiel bieten damit Sie es verwenden würden?

✂----------

6. Weiterer Studienverlauf

Möchten Sie über den weiteren Verlauf der Studie informiert werden? Die Antworten in diesem Abschnitt werden getrennt von den obigen verarbeitet und gespeichert.

6.1. □ Ich möchte über den weiteren Verlauf der Studie informiert werden.

6.2. □ Ich bin bereit eine im Rahmen der Studie entwickelte Smartphone-App zu testen.

6.3. E-Mail-Adresse _____

7. Dank

Vielen Dank für ihre Teilnahme an unserer Umfrage.

D Study Forms

The following pages contain the forms as we used them in the user study. We did not include all example URL but only one exemplary page with an attacked URL.

Beginnfragebogen

1. Generelles	1.1. Datum	1.2. Uhrzeit	1.3. Platznr.

2. Einschätzungen

Im Folgenden sollst du entscheiden wie sehr du den entsprechenden Aussagen zustimmst.

	1 Stimme gar nicht zu	2	3	4	5 Stimme voll zu
2.1. Ich kenne mich gut genug aus, um den Gefahren im Internet aus dem Weg zu gehen.					
2.2. Es fällt mir leicht legitime von gefälschten E-Mails zu unterscheiden.					
2.3. Es fällt mir leicht legitime von gefälschten Webseiten zu unterscheiden.					
2.4. Sicherheit im Internet bezieht sich ausschließlich auf Finanzanwendungen.					
2.5. Persönliche Daten gebe ich im Internet nur an Anbieter weiter, die ich gut kenne.					
2.6. Mir persönlich ist es egal, was mit meinen persönlichen Daten im Internet geschieht.					
2.7. Für den Schutz vor Gefahren im Internet ist jeder selbst verantwortlich.					
2.8. Ich vertraue E-Mails, die von mir bekannten Personen oder Organisationen kommen.					

Beispiel-Webseiten

zur Studie

„Internetsicherheit“

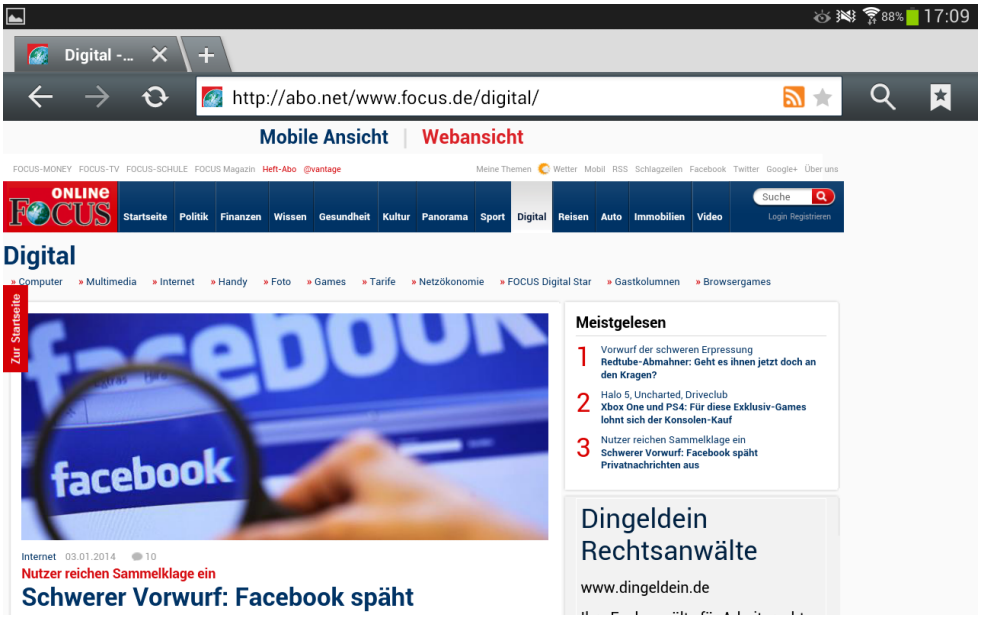
Datum:

Uhrzeit:

Sitzplatz:

OVorher ONachher

Bitte bearbeite diesen Fragebogen in der gegebenen Reihenfolge und blättere nicht zurück.



1. Würdest du auf dieser Webseite persönliche Daten eingeben?

O Ja O Nein

2. Umrande die Stelle, die dich am meisten zu deiner Antwort bewegt hat. Bitte markiere nur genau eine Stelle.

3. Wie sicher bist du dir?

Sehr unsicher O 1 O 2 O 3 O 4 O 5 sehr sicher

4. Kennst du „Focus“?

O Ja O Nein

5. Hast du bei diesem Anbieter ein Nutzerkonto?

O Ja O Nein

Abschlussfragebogen

1. Generelles

1.1. Datum	1.2. Uhrzeit	1.3. Platznr.

Zuerst ein paar generelle Fragen zu deiner Person.

1.4. Dein Alter....._____

1.5. Dein Geschlecht..... ☐Männlich ☐Weiblich ☐Anderes

1.6. Leidest du unter einer Farbenfehlsichtigkeit ? ☐Ja ☐Nein

1.7. Beruflicher Abschluss

☐Lehre ☐Meisterprüfung ☐Hochschulabschluss ☐Keiner ☐anderer: _____

1.8. In welchem Bereich arbeitest/studierst du zurzeit?....._____

1.9. Verwendest du ein Smartphone? ☐Nein ☐Ja, Android ☐Ja, iOS ☐Ja, _____

2. Benutzbarkeit

Im Folgenden sollst du entscheiden wie sehr du den entsprechenden Aussagen zustimmst. Es handelt sich hierbei um Standardfragen, daher kann es ggf. vorkommen, dass eine Aussage nicht ganz zutrifft.

	1 Stimme gar nicht zu	2	3	4	5 Stimme voll zu
2.1. Ich kann mir sehr gut vorstellen, die App regelmäßig zu nutzen.					
2.2. Ich empfinde die App als unnötig komplex.					
2.3. Ich empfinde die App als einfach zu nutzen.					
2.4. Ich denke, dass ich technischen Support brauchen würde, um die App zu nutzen.					
2.5. Ich finde, dass die verschiedenen Funktionen der App gut integriert sind.					
2.6. Ich finde, dass es in der App zu viele Inkonsistenzen gibt.					
2.7. Ich kann mir vorstellen, dass die meisten Leute die App schnell zu beherrschen lernen.					
2.8. Ich empfinde die Bedienung als sehr umständlich.					
2.9. Ich habe mich bei der Nutzung der App sehr sicher gefühlt.					
2.10. Ich musste eine Menge Dinge lernen, bevor ich mit der App arbeiten konnte.					

3. Meinung zur App

3.1. Der E-Mail Teil motiviert weiter zu machen.					
3.2. Die Menge der Texte war angemessen.					
3.3. Die Texte waren gut verständlich.					
3.4. Die App hat mir geholfen, Phishing Seiten zukünftig besser zu erkennen.					
3.5. Das Punktesystem mit den drei Leben pro Level war einfach zu verstehen.					

E Participant Recruitment for User Study

```
Sehr geehrte/r ....,

wir, Clemens Bergmann und Gamze Canova, sind Studenten des Fachbereichs Informatik an
   der TU Darmstadt und arbeiten zur Zeit unserer Masterthesis zum Thema
   Internetsicherheit.

Nun sind wir an dem Punkt angekommen, an dem wir eine Benutzerstudie (6.1.-11.1.2014)
   durchf\"{u}hren␣m\"{u}ssen, f\"{u}r␣die␣wir␣Teilnehmer␣suchen.␣Wir␣w\"{u}rden uns
   sehr freuen, wenn Sie uns hierbei unterst\"{u}tzen␣k\"{o}nnten. W\"{a}re␣es␣m\"{o}
   glich, den angeh\"{a}ngten␣Flyer␣mit␣dem␣unten␣stehenden␣Anschreiben,␣an␣Ihre␣
   StudentInnen␣und/oder␣Wissenschaftliche␣MitarbeiterInnen␣␣weiterzuleiten?␣Wir␣
   wissen,␣dass␣es␣schwierig␣sein␣k\"{o}nnte potentielle Teilnehmer kurz vor
   Weihnachten zu erreichen, jedoch k\"{o}nnen␣wir␣jede␣Unterst\"{u}tzung gebrauchen
   und w\"{u}rden␣uns␣\"{u}ber diese sehr freuen.

Wir bedanken uns im Voraus herzlich f\"{u}r␣Ihre␣Bem\"{u}hungen und w\"{u}nschen␣Ihnen
   ␣erholsame␣und␣besinnliche␣Festtage.

Mit␣freundlichen␣Gr\"{u}{\ss}en,
Clemens Bergmann und Gamze Canova

Text f\"{u}r␣Weiterleitung:
Im␣Rahmen␣unserer␣Masterarbeit␣haben␣wir␣eine␣Spiele-App␣entwickelt,␣die␣fachfremde␣
   Benutzer␣\"{u}ber Internetsicherheit informiert.  Diese App soll im Rahmen einer
   Benutzerstudie getestet werden. Hierzu brauchen wir deine Hilfe. Die Studie wird in
    Gruppen zu ca. 5 Personen in der zweiten Januarwoche (6.-10.) in Darmstadt
   stattfinden und insgesamt ca. 90 Minuten in Anspruch nehmen.  Der/Die Beste der
   Gruppe gewinnt einen Amazon-Gutschein im Wert von 20\euro.  Einzige Voraussetzung f
   \"{u}r␣die␣Teilnahme␣ist,␣dass␣du␣Erfahrung␣mit␣der␣Benutzung␣eines␣Smartphones␣
   hast.␣Bei␣Interesse␣oder␣Fragen␣erreicht␣ihr␣uns␣unter␣netstudy@cased.de.
```

Feindbild Internet?!

StudienteilnehmerInnen zur Internetsicherheit gesucht

Image courtesy of sheelamohan / FreeDigitalPhotos.net

TECHNISCHE UNIVERSITÄT DARMSTADT

SECUSO
SECURITY · USABILITY · SOCIETY

Im Rahmen unserer Masterarbeit haben wir eine Spiele-App entwickelt, die fachfremnde Benutzer über Internetsicherheit informiert.
Diese App soll im Rahmen einer Benutzerstudie getestet werden. Hierzu brauchen wir deine Hilfe.
Die Studie wird in Gruppen zu ca. 5 Personen in der zweiten Januarwoche (6.-10.) in Darmstadt stattfinden und insgesamt ca. 90 Minuten in Anspruch nehmen.
Der/Die Beste der Gruppe gewinnt einen amazon.de-Gutschein im Wert von 20€.
Einzige Vorraussetzung für die Teilnahme ist, dass du Erfahrung mit der Benutzung eines Smartphones hast.

Bei Interesse oder Fragen erreicht ihr uns unter netstudy@cased.de.

Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de
Internetsudie netstudy@cased.de

References

[1] A. Alnajim and M. Munro. An anti-phishing approach that uses training intervention for phishing websites detection. In *Information Technology: New Generations, 2009. ITNG '09. Sixth International Conference on*, pages 405–410, 2009.

[2] A. M. Alnajim. *Fighting internet fraud: anti-phishing effectiveness for phishing websites detection*. PhD thesis, Durham University, 2009.

[3] C. Amrutkar, P. Traynor, and P. Oorschot. Measuring ssl indicators on mobile browsers: Extended life, or end of the road? In D. Gollmann and F. Freiling, editors, *Information Security*, volume 7483 of *Lecture Notes in Computer Science*, pages 86–103. Springer Berlin Heidelberg, 2012.

[4] T. Amstad. *Wie verständlich sind unsere Zeitungen?* Abhandlung: Philosophische Fakultät I. Zürich. 1977. Studenten-Schreib-Service, 1978.

[5] Android. Android development - dashboards. `http://developer.android.com/about/dashboards/index.html`, 2013. Accessed: 2013-01-08.

[6] N. Arachchilage and M. Cole. Design a mobile game for home computer users to prevent from phishing attacks. In *Information Society (i-Society), 2011 International Conference on*, pages 485–489, 2011.

[7] N. A. G. Arachchilage, S. Love, and M. Scott. Designing a mobile game to teach conceptual knowledge of avoiding phishing attacksĂİ. *International Journal for e-Learning Security*, 2(2):127–132, 2012.

[8] T. Bakhshi, M. Papadaki, and S. Furnell. Social engineering: assessing vulnerabilities in practice. *Information management & computer security*, 17(1):53–63, 2009.

[9] I. Bank. The phishing game. `http://www.icicibank.com/online-safe-banking/Phishing-Game.html`. Accessed: 2013-01-17.

[10] A. Bergholz, J. De Beer, S. Glahn, M.-F. Moens, G. Paaß, and S. Strobel. New filtering approaches for phishing email. *Journal of computer security*, 18(1):7–35, 2010.

[11] M. Boodaei. Mobile users three times more vulnerable to phishing attacks. `http://www.trusteer.com/blog/mobile-users-three-times-more-vulnerable-to-phishing-attacks`, 2011. Accessed: 2013-12-26.

[12] Brightcove. Step-by-step guide to publishing in the android market on windows. `http://support.brightcove.com/en/app-cloud/docs/step-step-guide-publishing-android-market-windows`. Accessed: 2013-01-08.

[13] Brightcove. Step-by-step guide to publishing in the apple app store using a mac. `http://support.brightcove.com/en/app-cloud/docs/step-step-guide-publishing-apple-app-store-using-mac`. Accessed: 2013-01-08.

[14] J. Brooke. Sus: A retrospective. *Journal of Usability Studies*, 8(2):29–40, 2013.

[15] I. W. Business. Ranking: Die 100 gröSSten online-shops in deutschland 2011. `http://www.internetworld.de/Nachrichten/E-Commerce/Zahlen-Studien/Ranking-Die-100-groessten-Online-Shops-in-Deutschland-2011-Zalando-macht-zehn-Plaetze-gut-70245.html`, 2011. Accessed: 2013-01-24.

[16] M. Chandrasekaran, K. Narayanan, and S. Upadhyaya. Phishing email detection based on structural properties. In *NYS Cyber Security Conference*, pages 1–7, 2006.

[17] K.-T. Chen, J.-Y. Chen, C.-R. Huang, and C.-S. Chen. Fighting phishing with discriminative keypoint features. *Internet Computing, IEEE*, 13(3):56–63, 2009.

[18] T.-C. Chen, S. Dick, and J. Miller. Detecting visually similar web pages: Application to phishing detection. *ACM Trans. Internet Technol.*, 10(2):5:1–5:38, June 2010.

[19] S. Chiasson, M. Modi, and R. Biddle. Auction hero: The design of a game to learn and teach about computer security. In *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, volume 2011, pages 2201–2206, 2011.

[20] A. T. W. I. Company. Top sites in germany. `http://www.alexa.com/topsites/countries/DE`. Accessed: 2013-01-26.

[21] comScore Data Mine. Smartphones reach majority in all eu5 countries. `http://www.comscoredatamine.com/2013/03/smartphones-reach-majority-in-all-eu5-countries/`, 2013. Accessed: 2013-01-24.

[22] M. Csikszentmihalyi. Flow: The psychology of optimal performance, 1990.

[23] M. Csikszentmihalyi. *Finding flow: The psychology of engagement with everyday life*. Basic Books, 1997.

[24] D. Dealer. Data dealer. `https://datadealer.com/de`. Accessed: 2013-01-29.

[25] T. Denning, A. Lerner, A. Shostack, and T. Kohno. Control-alt-hack: The design and evaluation of a card game for computer security awareness and education. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, CCS '13, pages 915–928, New York, NY, USA, 2013. ACM.

[26] DIVSI - Deutsches Institut für Vertrauen und Sicherheit im Internet and Sozialwissenschaftliches Institut Nowak und Sörgel. *DIVSI-Milieu-Studie zu Vertrauen und Sicherheit im Internet: eine Grundlagenstudie*. Deutsches Institut für Vertrauen und Sicherheit im Internet (DIVSI), 2012.

[27] eBay. GefÃd'lschte e-mails erkennen und melden. `http://pages.ebay.de/help/account/reporting-spoof.html`. Accessed: 2013-01-11.

[28] I. Fette, N. Sadeh, and A. Tomasic. Learning to detect phishing emails. In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, pages 649–656, New York, NY, USA, 2007. ACM.

[29] G. Finance. Internet users by country. `http://www.gfmag.com/tools/global-database/ne-data/11942-internet-users.html#axzz2qr046GUb`, 2012. Accessed: 2013-01-19.

[30] R. Flesch. A new readability yardstick. *Journal of Applied Psychology*, 32(3):221–233, 1948.

[31] E. Gabrilovich and A. Gontmakher. The homograph attack. *Commun. ACM*, 45(2):128–, Feb. 2002.

[32] T. Goetz. Harnessing the power of feedback loops. *Wired Magazine*, 19(07), 2011.

[33] D. Goodin. From phishing to whaling. `http://www.theregister.co.uk/2008/04/16/whaling_expedition_continues/`, 2008. Accessed: 2014-01-19.

[34] GOV.UK. Guerilla testing. `https://www.gov.uk/service-manual/user-centered-design/user-research/guerilla-testing.html`, 2013. Accessed: 2013-01-24.

[35] A.-P. W. Group and C. CUPS. Phishing education landing page project. *Anti-Phishing Working Group*, 2009.

[36] A.-P. W. Group et al. Apwg global phishing survey. *Anti-Phishing Working Group*, 2013.

[37] A.-P. W. Group et al. Phishing activity trends report. *Anti-Phishing Working Group*, 2013.

[38] R. Gunning. *The Technique of Clear Writing*. McGraw-Hill, New York, 1952.

[39] A. I. Handbook. Us dept of transportation. *Federal Aviation Administration*, 2008.

[40] J. Hong. The state of phishing attacks. *Commun. ACM*, 55(1):74–81, Jan. 2012.

[41] IDC. Press release. http://www.idc.com/getdoc.jsp?containerId=prUS24442013, 2013. Accessed: 2013-01-24.

[42] M. Jakobsson and S. Myers. *Phishing and countermeasures: understanding the increasing problem of electronic identity theft*. Wiley. com, 2006.

[43] K. Jansson and R. von Solms. Simulating malicious emails to educate end users on-demand. In *Web Society (SWS), 2011 3rd Symposium on*, pages 74–80, 2011.

[44] K. Jansson and R. von Solms. Simulating malicious emails to educate end users on-demand. In *Web Society (SWS), 2011 3rd Symposium on*, pages 74–80, 2011.

[45] R. Koster. *Theory of fun for game design*. O'Reilly Media, Inc., 2010.

[46] P. Kumaraguru. *PhishGuru: a system for educating users about semantic attacks*. ProQuest, 2009.

[47] K. Lab. The evolution of phishing attacks: 2011-2013. *Kaspersky Lab*, 2013.

[48] E. Larkin. Spot the tiny phishing trick. http://www.pcworld.com/article/161232/tinyphish.html, 2009. Accessed: 2013-12-29.

[49] L. Lesbar. Testen sie ihren text. http://leichtlesbar.ch/html/. Accessed: 2013-01-28.

[50] W. Liu, X. Deng, G. Huang, and A. Fu. An antiphishing strategy based on visual similarity assessment. *Internet Computing, IEEE*, 10(2):58–65, 2006.

[51] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker. Beyond blacklists: Learning to detect malicious web sites from suspicious urls. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '09, pages 1245–1254, New York, NY, USA, 2009. ACM.

[52] S. Marchal, J. François, R. State, and T. Engel. Proactive discovery of phishing related domain names. In D. Balzarotti, S. J. Stolfo, and M. Cova, editors, *Research in Attacks, Intrusions, and Defenses*, volume 7462 of *Lecture Notes in Computer Science*, pages 190–209. Springer Berlin Heidelberg, 2012.

[53] McAfee. The economic impact of cybercrime and cyber espionage. *Center for Strategic and International Studies*, 2013.

[54] A. McMahan. Immersion, engagement and presence. *The video game theory reader*, pages 67–86, 2003.

[55] Microsoft. How to recognize phishing email messages, links, or phone calls. http://www.microsoft.com/security/online-privacy/phishing-symptoms.aspx. Accessed: 2013-01-11.

[56] Microsoft. Identify fraudulent e-mail and phishing schemes. http://office.microsoft.com/en-001/outlook-help/identify-fraudulent-e-mail-and-phishing-schemes-HA001140002.aspx. Accessed: 2013-01-11.

[57] T. Moore and R. Clayton. Examining the impact of website take-down on phishing. In *Proceedings of the Anti-phishing Working Groups 2Nd Annual eCrime Researchers Summit*, eCrime '07, pages 1–13, New York, NY, USA, 2007. ACM.

[58] T. Moore and R. Clayton. How hard can it be to measure phishing? *Mapping and Measuring Cybercrime*, 2010.

[59] C. Murphy. Why games work and the science of learning. In *Interservice, Interagency Training, Simulations, and Education Conference*, 2011.

[60] A. Obied and R. Alhajj. Fraudulent and malicious sites on the web. *Applied Intelligence*, 30(2):112–120, 2009.

[61] C. K. Olivo, A. O. Santin, and L. S. Oliveira. Obtaining the threat model for e-mail phishing. *Applied Soft Computing*, 2011.

[62] OnGuarOnline.gov. Invasion of the wireless hackers. http://www.onguardonline.gov/media/game-0006-invasion-wireless-hackers. Accessed: 2013-01-17.

[63] OnGuarOnline.gov. Mission laptop security. http://www.onguardonline.gov/media/game-0008-mission-laptop-security. Accessed: 2013-01-17.

[64] OnGuarOnline.gov. Phishing scams (game). http://www.onguardonline.gov/media/game-0011-phishing-scams. Accessed: 2013-01-11.

[65] PayPal. Was ist phishing? https://www.paypal.com/de/webapps/mpp/phishing. Accessed: 2013-01-11.

[66] Phishing.org. Phishing techniques. http://www.phishing.org/phishing-techniques/. Accessed: 2013-01-19.

[67] PhishTank. Phishtank. http://www.phishtank.com/, 2013. Accessed: 2013-12-29.

[68] P. Prakash, M. Kumar, R. Kompella, and M. Gupta. Phishnet: Predictive blacklisting to detect phishing attacks. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–5, 2010.

[69] S. Purkait. Phishing counter measures and their effectiveness–literature review. *Information Management & Computer Security*, 20(5):382–420, 2012.

[70] Z. Ramzan. *Phishing Attacks and Countermeasures*. Springer Berlin Heidelberg, 2010.

[71] RSA and ECM. Phishing kits - the same wolf, just a different sheep's clothing. *Fraud report*, 2013.

[72] N. M. Sadeh. Why phish should not be treated as spam. http://www.drdobbs.com/security/why-phish-should-not-be-treated-as-spam/240001777, 2012. Accessed: 2013-01-19.

[73] G. C. Safe et al. Phishing: How many take the bait? http://www.getcybersafe.gc.ca/cnt/rsrcs/nfgrphcs/nfgrphcs-2012-10-11-eng.aspx, 2012. Accessed: 2013-12-29.

[74] J. Schell. *The Art of Game Design: A book of lenses*. Taylor & Francis US, 2008.

[75] P. Schöll. Flesch-index berechnen. http://www.fleschindex.de. Accessed: 2013-01-28.

[76] B. Schwartz. *The paradox of choice*. HarperCollins, 2009.

[77] R. secure. Phishing for disaster: the cost of corporate ignorance. *Whitepaper about the effects of corporate ignorance of phishing*, 2010.

[78] M. E. Seligman. *Flourish: A visionary new understanding of happiness and well-being*. Simon and Schuster, 2012.

[79] 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*, SOUPS '07, pages 88–99, New York, NY, USA, 2007. ACM.

[80] D. P. Simon. The art of guerilla usability testing. http://www.uxbooth.com/articles/the-art-of-guerilla-usability-testing/, 2013. Accessed: 2013-01-24.

[81] SonicWALL. Sonicwall phishing iq test. http://www.sonicwall.com/furl/phishing/. Accessed: 2013-01-11.

[82] SPAMfighter. Becoming more difficult to detect phishing email attack, says security experts. http://www.spamfighter.com/News-18495-Becoming-More-Difficult-to-Detect-Phishing-Email-Attack-says-Security-Experts.htm, 2013. Accessed: 2013-01-11.

[83] Stilversprechend. Stilversprechend. http://stilversprechend.de/index.html. Accessed: 2013-01-28.

[84] Symantec. Race to stay safe. https://www.staysecureonline.com/staying-safe-online/. Accessed: 2013-01-11.

[85] T. A. R. Team et al. Spear-phishing email: Most favored apt attack bait. http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-spear-phishing-email-most-favored-apt-attack-bait.pdf, 2012. Accessed: 2013-12-29.

[86] W. S. Technologies. Anti-phishing phyllis. http://www.wombatsecurity.com/antiphishingphyllis. Accessed: 2013-01-11.

[87] E. L. Thorndike. *The fundamentals of learning*. Teachers College Bureau of Publications, 1932.

[88] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6):80–83, 1945.

[89] H. Zhang, G. Liu, T. W. S. Chow, and W. Liu. Textual and visual content-based anti-phishing: A bayesian approach. *Neural Networks, IEEE Transactions on*, 22(10):1532–1546, 2011.

[90] J. Zhang, P. A. Porras, and J. Ullrich. Highly predictive blacklisting. In *USENIX Security Symposium*, pages 107–122, 2008.

[91] W. zu Wem Firmenverzeichnis. Top 100 banken in deutschland. http://www.wer-zu-wem.de/ranking/banken/, 2011. Accessed: 2013-01-24.