Anti-Phishing Education App

Design, Implementation and Evaluation

Master-Thesis von Clemens Bergmann und Gamze Canova

Januar 2014

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Fachbereich Informatik
Security, Usability and Society

Anti-Phishing Education App
Design, Implementation and Evaluation

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 16, 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 16, 2014

(G. Canova)

Contents

Abstract

...

1 Introduction

This chapter introduces the target of this work, which is to design, implement and evaluate an educational app. The app is supposed to help unexperienced users to detect phishing attacks. At first we are going to motivate the benefit of our work and how we envision our approach to achieve our goal. Next, we define our specific objectives and finally, we provide an overview of the following chapters.

1.1 Motivation

Nowadays, a world without Internet is unimaginable for many people. However, it is undeniable that with the great benefits of the internet also come great threats. One major issue of today's digitalized world is spam in general and phishing in detail. Phishing is a for of Fraud to lure confidential information from users, cf. Section 2.1. The goal of the Attacker is to impersonate the User in online systems. This can be used to access corporate systems, damage the users reputation or simply steal money from him. Usually, phishing happens through fake websites which imitate the original ones. On these so called phishing websites the users are asked to enter their personal or account data. In this section we elaborate on the importance of countering such phishing attacks with the aid of user education.

1.1.1 Consequences of Phishing

Falling for a phishing attack has several consequences for the fooled person as well as for the target company or organization. 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. 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 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. Moreover, the targeted institutions may sustain a damaged reputation due to phishing attacks. 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 will tell other people about this displeasant experience. Ultimately, these victims will lose their trust in the institution targeted by the phisher. Moreover, they might lose confidence in eCommerce operations and the Internet in general.

1.1.2 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 [**?**]. 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 attacks, 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 [**?**]. 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 [**?**]. Finally, [**?**] estimates worldwide costs caused by phishing at about $1.5 billion for the year 2012.

1.1.3 Technical Solutions to Counter Phishing

Several technical solutions to counter phishing have already been suggested in literature [**?**]. Unfortunately, those techniques did not seem to suffice, which is also reflected by several phishing stastics, cf. Section 1.1.1. In the following some of these approaches are briefly summarized.

Spam filters Commonly, the phisher sends out a tremendous volume of emails to random users which contain links to fake websites. There the users are lured to disclose their personal data. Consequently, one possible countermeasure to stop phishing is to filter these e-mails before they even reach the receiver. Various approaches for such spam filters

already exist [?, ?, ?], but spam filters also have their drawbacks. First, a spam filter needs to be installed and applied to the users' e-mail accounts. When using a E-Mail Service the service provider is in general not allowed to access the users' mail without his permission. Therefore spam filters only apply to users that actively enable them. Second, even if spam filters are used by the majority, one can not make sure that they are updated regurlarly. Moreover, phishers are constantly improving their techniques to circumvent current spam filters. Consequently, such filters can not assure 100% accuracy. 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 [?].

Blacklists Fake websites are a common way for phishers to get at users' data. Thus, another alternative to protect potentially endagered users from phishing attacks is to restrict the access to such 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 [?, ?]. 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 [?, ?]. Nevertheless, there is no flawless blacklisting approach, as there are always malicious websites which can bypass such protective systems. Moreover, these systems require a high effort to maintain, since a regular and realtime update is inevitable in order to make the system effective [?]. Furthermore, there is the weakest link in the security chain: the users who are very often unsure about what to do when getting such security warnings [?]. As a matter of fact, in case of disregard of these warnings such systems are useless.

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 malicously duplicated websites mainly based on visual similarities [?]. Various solutions can be found in literature to approach this [?, ?, ?]. 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. 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 as always the human factor plays a major role here: if users keep misunderstanding or ignoring the provided visual indicators such techniques will remain of no use.

Takedown Commonly, hosting providers are urged to take down revealed malicious websites by certain parties: banks, other organizations or specialized takedown companies, for example [?]. 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. However, this approach - similar to the blacklist one - cannot defeat phishing entirely, since it is not fast enough [?]. During the uptime of the fraudulent website falling for it remains a threat.

In conclusion, there are two major issues of existing techniques. First, 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. The second major problem with these approaches is the user behavior. As already indicated above users tend to overlook or intendedly ignore security warnings. If the user behavior does not change such approaches will in fact remain unhelpful. The problem here 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 or they just implicitly assume the system to be secure. Another factor for overlooking and ignoring these warnings is the lack of security awareness of the users. 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. In summary, pure technical solutions cannot guarantee 100% protection. On the other hand, in the end the users themselves fall into the traps of the attackers. 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 educating them how to protect themselves against such traps.

1.1.4 Anti-Phishing Education on the Smartphone

There are several reasons why we chose to educate users on the smartphone. The main characteristic of a smartphone is that it is mobile and enormously smaller than the well-known desktop computers. As a consequence there is much

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. 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 [**?**], 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 mail clients effectively provide no way to check the validity of a incoming mail. In addition, given the fact that the majority of the people use a smartphone on a regular basis **add to refs(http://www.comscoredatamine.com/2013/03/smartphones-reach-majority-in-all-eu5-countries/)**, there is a need for the protection of smartphone users. 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, so that probably more users would be willing to use it. Second, it is easy to transfer the knowledge of smartphones to desktop computers as the screen is bigger and the URL is easy to find. Transferring knowledge from desktop computers to smartphones, on the other hand, is not that simple.

1.2 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 educate the user 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 user awareness seems to be a major issue concerning the security-related user behavior. For this reason we want to raise the users' 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 links do not necessarily lead to the target website the link displays to the user. 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, the user should practically experience these aspects and not only be told, since being told will not suffice to motivate the user to go on with the app. Moreover, besides technical solutions, valuable information has to be made available to the user. In particular, we want to qualify our app users to detect phishing URLs so they can distinguish phishing websites from legitimate ones. For the app we focus on educating the user. Increasing the security awareness is a minor introductory part of the app.

1.3 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 begin with the definition of phishing and proceed with distinguishing several phishing techniques, attack channels and phishing variations.

2.1 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.

"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." [**?**]

2.2 Phishing Techniques

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

Deceptive Phishing In deceptive phishing social engineering plays a key role. Here, users are lured to disclose 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 Based Phishing 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 Phishing 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 Phishing 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. Cross-site scripting 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.

2.3 Phishing Attack Channels

These days several attack channels exist that can be used by phishers to reach their victims. This section introduces possible attack channels [**?**].

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.

2.4 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. 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 [**?**].

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. 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[**?**, **?**].

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.

In the following section we will summarize the scope of the term phishing for this work.

2.5 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 [**?**, **?**, **?**]. 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 (Mail-Adress, 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.

3 Related Work

This chapter deals with previous work on anti-phishing education. We divided the related work we have found in literature into two categories: the *content*, i.e. what the user is taught, and the *medium*, i.e. how the user is taught. In the following, we will provide an overview of this content and medium classification. Subsequently, we will provide examples of previous work.

3.1 Content Classification

The objective of this section is to introduce the different classes of learning content which we could identify in previous work. **ICH BIN MIR UNSICHER OB MELANIE DIESE KLASSIFIZIERUNG GEFIEL. VIELLEICHT SOLLTEN WIR TATSCÃĎHLCIH EINFACH NUR RUNTERZÃĎHLEN WAS ES SO GIBT UND AUF VOR UND NACHTEILE EINGEHEN**

General Knowledge Transfer Renowned and targeted websites, such as PayPal, eBay or Microsoft provide general and superficial information about phishing [**?**, **?**, **?**]. 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.

E-Mail Based Knowledge In this class of content, the user is told about the "anatomy" of phishing e-mails [**?**, **?**]. 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 unknownlingly 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 [**?**, **?**]. Even if today still many phishing e-mails exhibit the obvious characteristic of having no personal salutation or being urgent and threatening, 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 then, 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 [**?**, **?**]. Links to phishing websites are not only distributed by phishing e-mails. Such links can be spread via any communication channel. It is even possible to land on a phishing website by just surfing. Thus, for these cases knowing how to determine whether an e-mail is fake or legitimate is of no use. In these situations knowing how to distinguish phishing URLs from valid ones will help. The problem with this approach is that as soon as the DNS or host file is attacked, cf. Section 2.2, even for experts it will get difficult to distinguish a phishing website from the legitimate one.

3.2 Medium Classification

The objective of this section is to introduce the different classes of learning media which we could identify in previous work.

Game Based Learning One way to communicate the learning content to the user is to use the traditional game. Such a game usually has a "background story" and a "mission" the user has to accomplish [**?**, **?**]. The game design is important and depends on the target group. Previous work, for example, has focused on a fish as starring role in their game, cf. Section 3.3. 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 prestudy, cf. Section 6.

Quiz Based Learning The quiz based approach is a form of a game which relies on a question-answer cycle without using a specific background story [**?**]. 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 [**?**]. 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, since they look almost identical. However, 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 [**?**, **?**]. 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. However, the missing positive feedback is a major flaw of this strategy. The user is only notified in case of a mistake and not in case he has successfully rejected to react to the simulated phishing e-mail. A further problem is raised with the realization of such an approach. Legal issues will arise when sending simulated phishing e-mails which claim to com from a reputable vendor, for example, Amazon.

3.3 Previous Work

Previous work here ... (e.g. Anti-Phishing Phil and Phyllis)

4 Focus

Introductory sentences... Based on the discussion of the previous section we decided to..... (not only based on previous section)

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 We focus on the detection of spoofed websites resp. 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. However, they can also be spread via SMS or even phone calls. Ultimately, a phishing website can also be reached by just surfing in the Internet. As a consequence, our approach covers all attack channels, as long as the user is tricked to divulging sensitive information via a phishing website.

Mass Phishing as Variation of Phishing We cover in particular mass phishing. 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 **wenn oben geschrieben, hier auch schreiben**

URL Based Knowledge as Learning Content The advantages of telling the user what to pay attention to within e-mails are the following: if the user recognizes the phishing e-mail before clicking on a link he does not even get onto a fake website where he could be lured to divulge his credentials. This also would mean, that the user would not be forwarded to a page where a malicious download might be started. On the other hand, these fake e-mails become more and more sophisticated and thus it becomes harder to distinguish them from legitimate ones [**?**, **?**]. Additionally, e-mail is not the only attack channel where links to phishing websites can be distributed. Those links are also spread via instant messaging systems, online social networks, or SMS, where the messages would differ from those in e-mails. Moreover, phishing websites can also be reached by surfing [**?**], where the e-mail based knowledge approach would completely fail. For these reasons we decided to focus on communicating URL based knowledge to the user. This way, the disadvantages of e-mail based knowledge are mitigated. Furthermore, we believe that URL based knowledge gives the most reliable hint regarding its "belonging", i.e. whether a URL in fact belongs to a legitimate website or not.

"After Click" URL Analysis We have decided to consider the "after click" scenario for the following reasons: Firstly, we cannot hinder users from clicking on links and make them type in the whole URL into the address bar. This is too effortful, especially on smartphones, and thus will not be followed by them. Secondly, many links contain redirects. Such

redirects are not recognizable before the click. A further problem the "before click" scenario raises is that the stock e-mail client of Android does not provide the functionality of viewing the destination URL before clicking on it. The only way to have a look at the URL before clicking on it is to make a long press onto the link, copy it into the clipboard, paste it somewhere else and then analyze it. Then, after the analysis the URL has to be sent to the browser. However, as this is also involves too much effort, no user will follow such a suggestion. Finally, even if there are many other e-mail clients which offer viewing the destination URL via long press only, we believe that this should not be communicated to the user for two reasons. Firstly, we do not know how many Android users actually make use of the stock e-mail client, which does not offer this functionality. Secondly, and most importantly, this functionality has the potential to mislead the user. A drawback of the URL destination preview is that the end of it 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 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 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. For the reasons explained above we have decided to consider the "after click" scenario. This approach suffers the disadvantage that users might click on a link which has a download of malware behind it. For now, we consider this as future work, as there is no possibility to detect whether there is a download behind a URL before requesting the site.

Considered Browser fÃijr screenshots benutzen wir Android standard browser, aber kann auf jeden browser Ãijbertragen werden. wir haben Ãijberlegt, sahcen wie lock icon Blabla einzubauen, aber da immer sehr unterschiedlich haben wir uns dagegen entschieden. um unsere Methode mÃűglichst allg. zu halten– siehe bitte section blablabla

4.2 System Requirements

In the following we are listing the system requirements which need to be met for the final app.

Android We have decided to develop an app for the Android operation system as we believe we have greater freedom here compared to an iOS app. The publication of an iOS app, for example, is connected with more requirements, which is not the case for Android apps [**?**, **?**].

Version Our original 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 []. For this reason we have decided to modify the code so that these users can also install and use our app.

Android Standard Browser Android standard browser is kein system requirement - raus. muss irgendwoanders erwÃđhnt werden. (was betrachten wir beim erklÃđren oder so)...

4.3 Assumptions

Secure DNS ...

Secure Smartphone ...

No Before-Click URL Analysis ...

Download URLs Possible ...

4.4 Limitations of Our Approach

Cross-Site Scripting ...

URL Hiding Techniques ...

5 Target Group

Introductory sentences... DIVSI

6 Pre-Survey

Before elaborating on the concrete app design we ran a small pre-survey. To the best of our knowledge there do not exist other surveys which resemble ours and additionally were conducted in Germany. This chapter deals with the main objectives of the pre-survey. Furthermore, it provides some details and finally presents the results and evaluates the questionnaire.

6.1 Main Objectives

Our main objectives of this pre-survey were twofold:

1. **Awareness and Knowledge** One goal of the pre-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 on knowledge on the topic of 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 some 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. **SURVEY IN APPENDIX???**

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 regaring 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 **..ADD REF**

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 about and think of 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, or whether he finds a text-based approach meaningful as well as whether he would have fun with a question-answer quiz game.

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

6.2.2 Distribution

In total 251 persons participated in our pre-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 pre-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 parents whether they can ask their friends, colleague 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 so we could reach a wider range of people.

6.2.3 Filtering for Evaluation

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

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 filter 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 filter out 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 filter out any participant with this question.
User reads up on topics related to dangers in the Internet	Participants who have chosen "no" as answer are filtered out. We specifically target users who are interested in getting safer in 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 any participant with these statements.
Section to questions concretely related to phishing	We do not filter out any participant with these statements.
Section to preferences for an anti-phishing education app	We do not filter out any participant with these statements.

In the following section we present and evaluate the results of the pre-study. With the filtering above we had 169 remaining participants who were considered for the evaluation.

6.3 Results and Evaluation

The pre-studey yielded interesting results which should be considered when designing an anti-phishing education app, either for this work, or if not possible due to time constraints in future work. This section outlines the results of the pre-study.

General Information The ratio of our male and female 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 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 1 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 depends on the expertise of the person being online. However, the more often a user is in the Internet, the more likely it is that he will experience a phishing attempt.

Usage Distribution of Internet Applications Figures 2 and 3 summarize the usage distribution of Internert 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, cf. Section 2.3, e-mail is a common attack channel for phishing attempts. Consequently, all of these e-mail application users are potentially endangered by phishing attacks. The same applies to participants using browsers. A common way to trick users to disclose their confidential information is the use of fake websites. These websites can be reached by clicking on a link in an e-mail, SMS or in online social networks or instant messaging systems as well as by simply surfing in the Internet. About 80% of all considered participants make use of desktop or smartphone browsers. Furthermore, it is conspicuous that banking is far less used on the smartphones compared to desktop computers. While about 74.56% of the participants make use of online banking on the desktop computer, only 26.63% make use of it on their smartphones, which is however still a quarter of the participants. The question to ask here is if these users use the browser for the online banking or if they use apps provided by their bank. Regardless of the answer to this question, these users might be more likely to react to phishing e-mails, claiming to come from their bank, on their smartphone 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 1.1.4. To sum it up, all the 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 pre-study), depending on the usage distribution.

Self-Assessment - Knowledge to avoid dangers of Internet With a Likert scale the participants had to indicate how much they agreed with the following statement: “I have enough knowledge to avoid the dangers of the Internet”. 18.34% of the participants strongly agree with this self-assessing statement. 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 With a Likert scale the participants had to indicate how much they agreed with the following statement: “I find it easy to distinguish legitimate e-mails from fake ones”. Here, 37.23% of the participators strongly agreed with the statement and even 50% of them agreed with it. Only about 8% of the participant 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 to e-mails from known parties With a Likert scale the participants had to indicate how much they agreed with the following statement: “I trust e-mails which come from persons I know”. 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 agains shows, that most of the participants are not aware that spoofing the “from field” of an e-mail is very easy. These users are likely to react to e-mails which claim to be, for example, from friends. Such e-mails may 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 in 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, originally, 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, if he had fallen

for such an attack in reality. 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, so it is something which should be considered in future work.

Services endangered by phishing Figure 4 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 with fake websites (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 in the Internet, however they are potential targets and should be communicated to the user, for example, with the aid of the choice of the URLs to decide on.

Data endangered by phishing Figure 5 outlines the results of this question and illustrates that the participants agree that every kind of data is endangered by phishing attacks. 90.53% of the participants are of the opinion that login data is endangered by phishing. About 89% agree that credit card information as well as personal data is endangered, too. Finally, 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 todo. evtl nochmal aufteilen

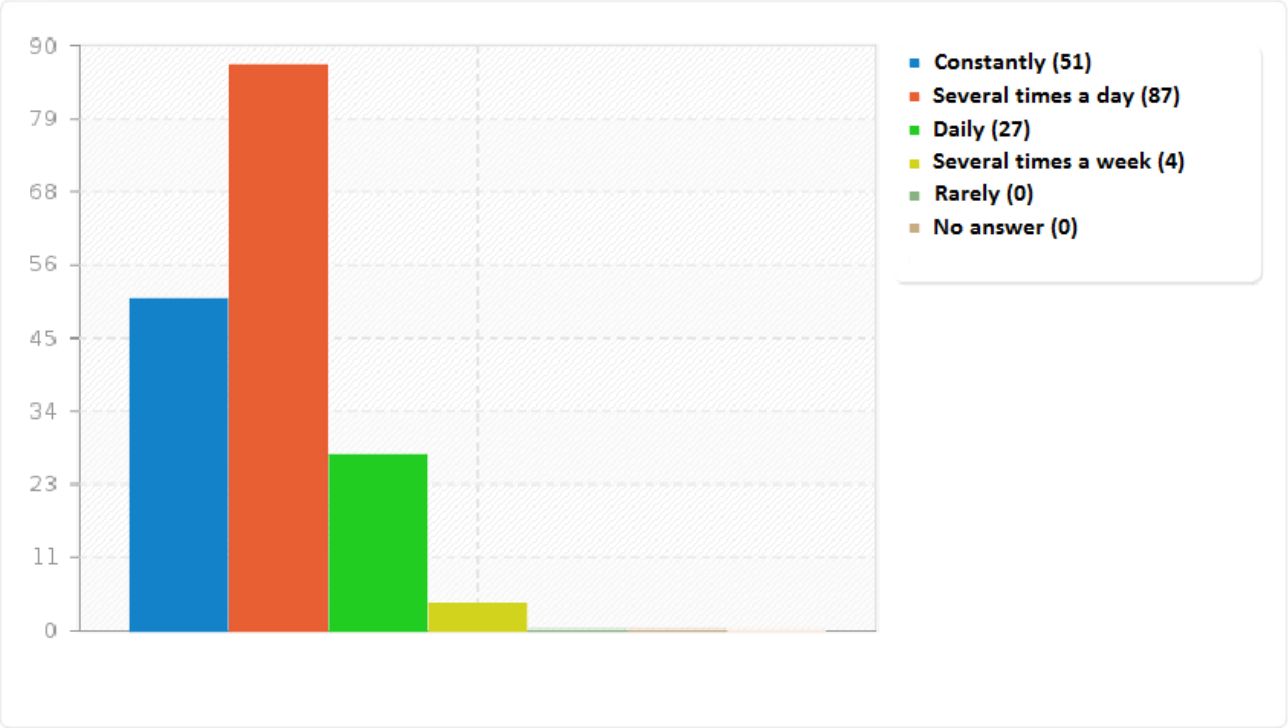

Figure 1: Frequency of Internet Usage

7 Teaching and Learning Content

TEILE VON DIESEM KAPITEL SCHEINEN NACH MELANIE EHER BACKGROUND ZU SEIN: SCHAUEN In this section we will describe and elaborate on different teaching and learning contents which can potentially be communicated to the user. At the same time we will reason our decision whether to communicate the specific content or not.

7.1 Phishing URLs

As aforementioned, we focus on teaching the user how to analyze a given URL and to decide on it 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 how the URLs are spoofed in order to deceive the users. For the analysis of phishing

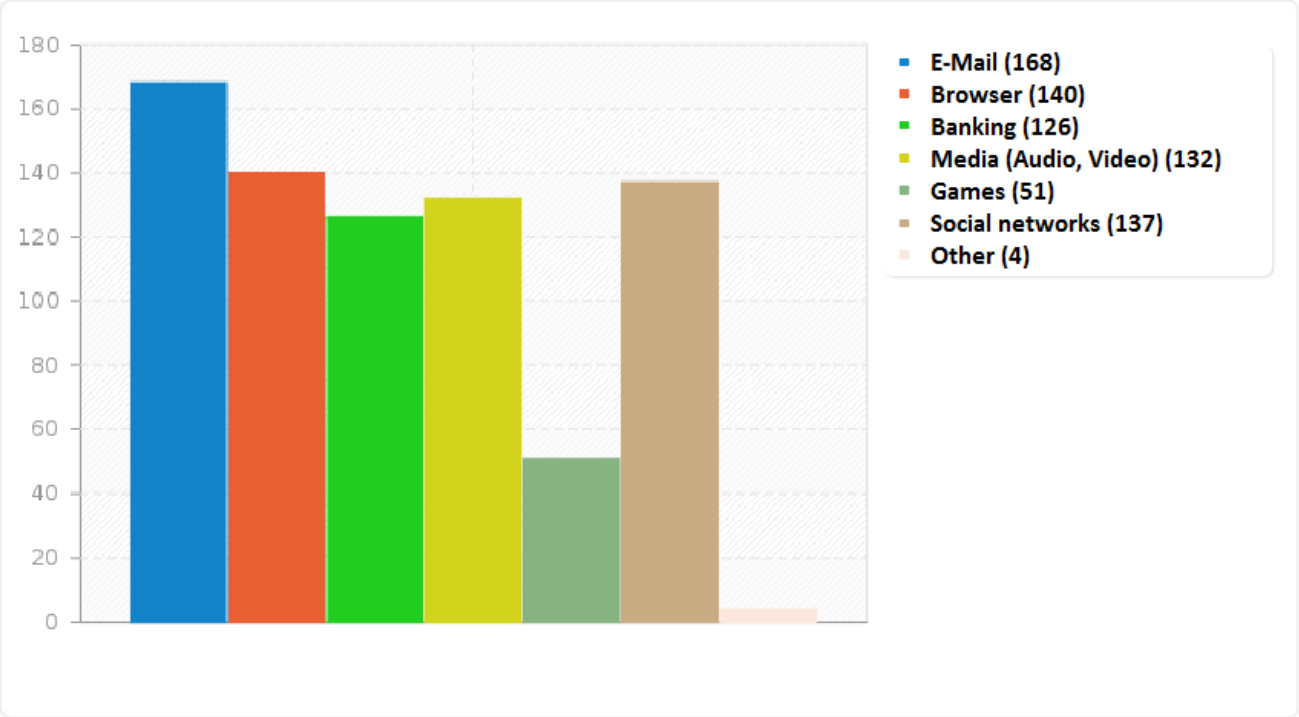

Figure 2: Usage of Internet Applications on Desktop Computers

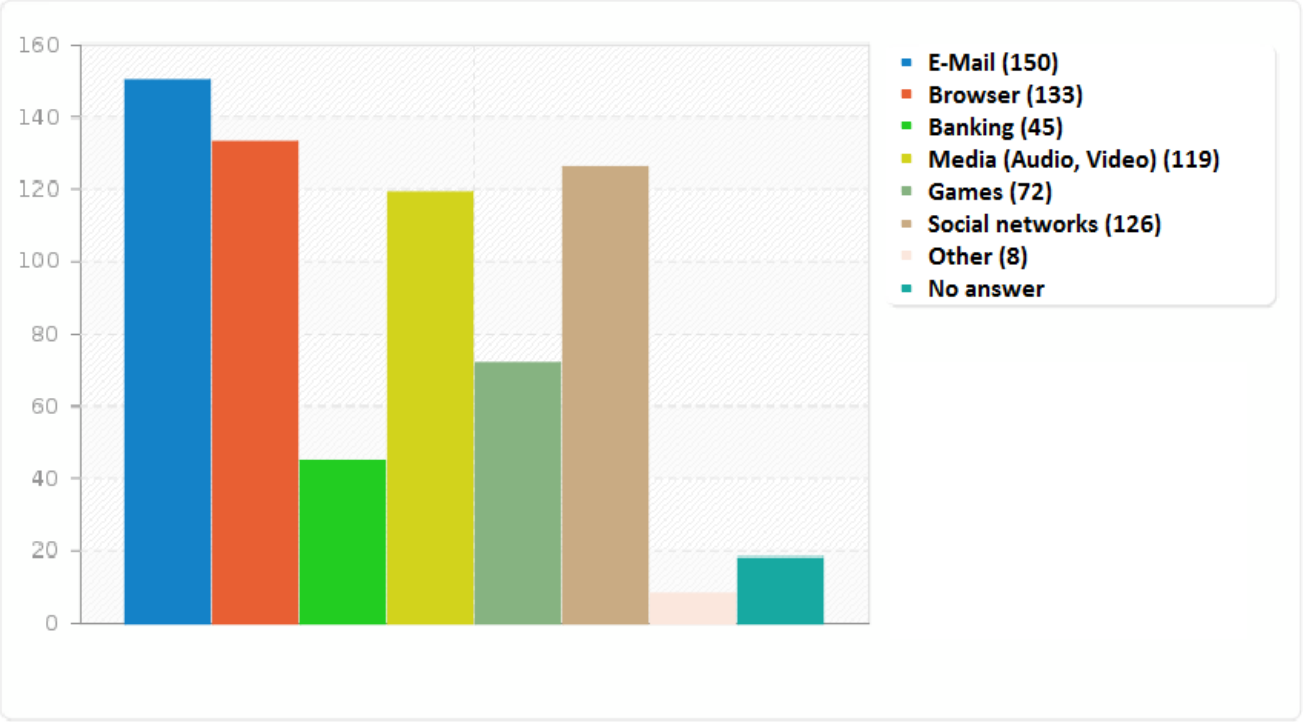

Figure 3: Usage of Internet Applications on Smartphones

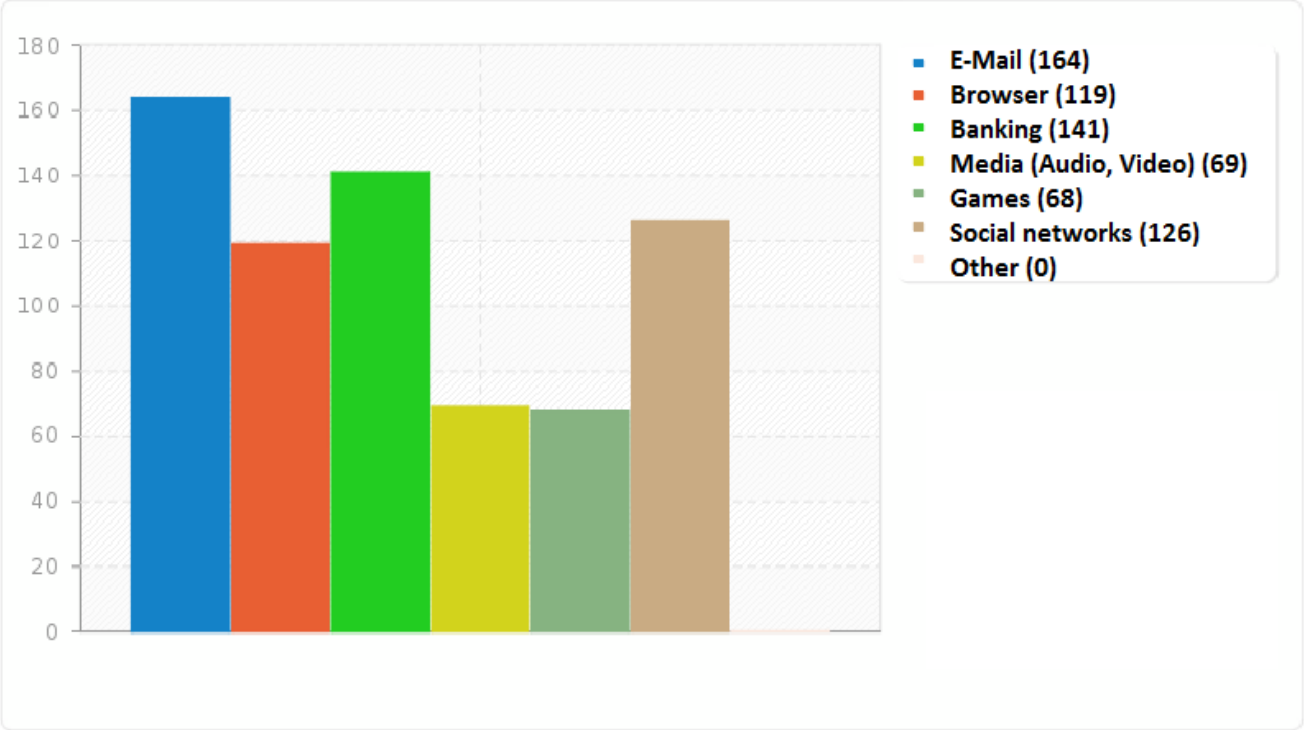

Figure 4: Services Endangered By Phishing

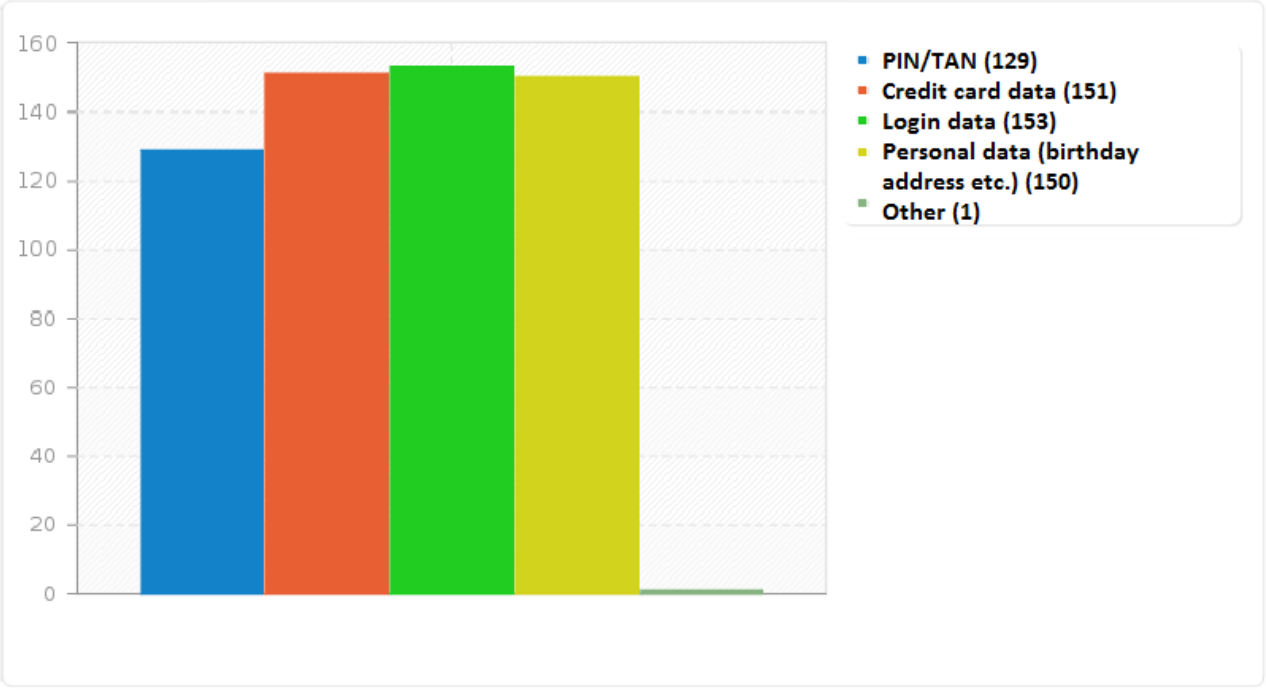

Figure 5: Data Endangered By Phishing

URLs we chose the database of PhishTank. 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. Renowned organizations such as Yahoo, Kaspersky Lab and Mcafee use the data submitted by PhishTank [**?**]. A further deciding reason to choose PhishTank as our phishing URL database was that Kaspersky Lab itself recommended us to make use of it for our URL analysis. For the phishing URL analysis we made use of the URL categories which had been identified by the authors of Anti-Phishing Phil [**?**] as a starting point. To these belong IP address URLs, subdomain URLs as well as similar and deceptive domain URLs. With these given categories we tried to assign the PhishTank URLs to the available 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.

7.1.1 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 about the authenticity of an URL when it contains the brand name anywhere. Phishers exploit this lack of knowledge in different way. In the following we present the identified spoofing attacks on URLs and state whether they are covered by the app.

Subdomain Phishers make use of subdomains which are very similar or even identical to the domains of the spoofed target institutions. This makes the users believe that they are on a legitimate website. This form of URL spoofing is covered by our education app.

IP Address Sometimes phishers do not even bother registering any domain at all. In this case, the URL to the phisher's fake website contains an IP address. This form of URL spoofing is covered by our education app.

Nonsense Domain We frequently encountered URLs which had registered quite nonsense as their domain. The domain names ranged from random letters to domain names like "marketstreetchippy.com". Sometimes other parts of the URL contained the brand name, but sometimes there was no clue in the URL about to where it is actually leading. This form of URL spoofing is covered by our education app.

Trustworthy, But Unrelated Domain Some URLs are very well-crafted. When reading them they appear meaningful and trustworthy. This is particularly accomplished by making use of domain names which sound very trustworthy, 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 can be perfectly deceived. This form of URL spoofing is covered by our education app.

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 has 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. This form of URL spoofing is covered by our education app.

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 langauge sets, such as Cryllic languages, where the characters will look almost identical [**?**]. The letter case is indistinguishable for the human eye in many cases. 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 URL 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. 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 quite popular [**?**]. Tiny URLs redirect the tiny URL to the actual long URL. 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.

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"[**?**]. 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.

7.1.2 Problems and Challenges With The Categorization

7.2 Android Elements

Eventuell Titel umbenennen, anders strukturieren...

Invisible Address Bar Find URL Bar, Browser

Use of Https Within Websites Browser

Analyze Complete URL Via Address Bar Browser

Show URL Before Click In E-Mail (not always possible), while surfing (long touch)

Copy and Paste URL too much effort, additionally: redirects still possible

7.3 Android Browser Security Indicators

Https Padlock Browser

Displayed Webaddress on Https Sites Browser

Certificate Verification

Touch Padlock to see whole URL.. problems: see document...

7.4 E-Mail Spoofing

From Field not trustworthy

E-Mail Content in hand of attacker

Links in E-Mails do not necessarily go where it claims to go (not only in e-mail links).

7.5 General Recommended Behavior

Do Not Click

Do Not Download Attachment

Look at URL

Data Economy

Date Entry Via Https

7.6 Conclusion / Summary

Summarize what to communicate to user here...

8 Approach for Our Anti-Phishing Education App

This chapter presents our final approach for the Anti-Phishing Education 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, in this particular case, against phishing.
    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 free text. Upon submitting the form the user will receive an e-mail with the e-mail address he had indicated as sender. The free 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. The user will learn that he cannot fully trust the "from" field and the content of the e-mails he is receiving.
    b) **Linktext 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. He is also told that this app focuses on exactly this.
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) **Information Material** The second part of the app is divided into levels of increasing difficulty. Here 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. This is done in a very simplified way, so that even unexperienced users can 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 can be consulted in Section 8.4.
    b) **Exercise to Information Material** After every introductory material in each level an exercise section is followed. 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 second- and top-level domain" information material the user gets a couple of valid URLs of which he has to identify the second- and top-level domains. 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 second- and top-level domain. Only if the user correctly identifies the second- and top-level domain he receives the points for this URL, otherwise the answer to this URL is considered wrong, because we assume that the user has just guessed in this case.
    c) **Repeat 2.a and 2.b With Increasing Difficulty** There is an increase of difficulty in each level. That is to say, in each level it gets more difficult to distinguish phishing URLs from valid ones, cf. Section 8.4.

The next section deals with the concrete rules of the game.

8.2 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. 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. 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. The following Figure 6 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.

The following section deals with our leveling strategy.

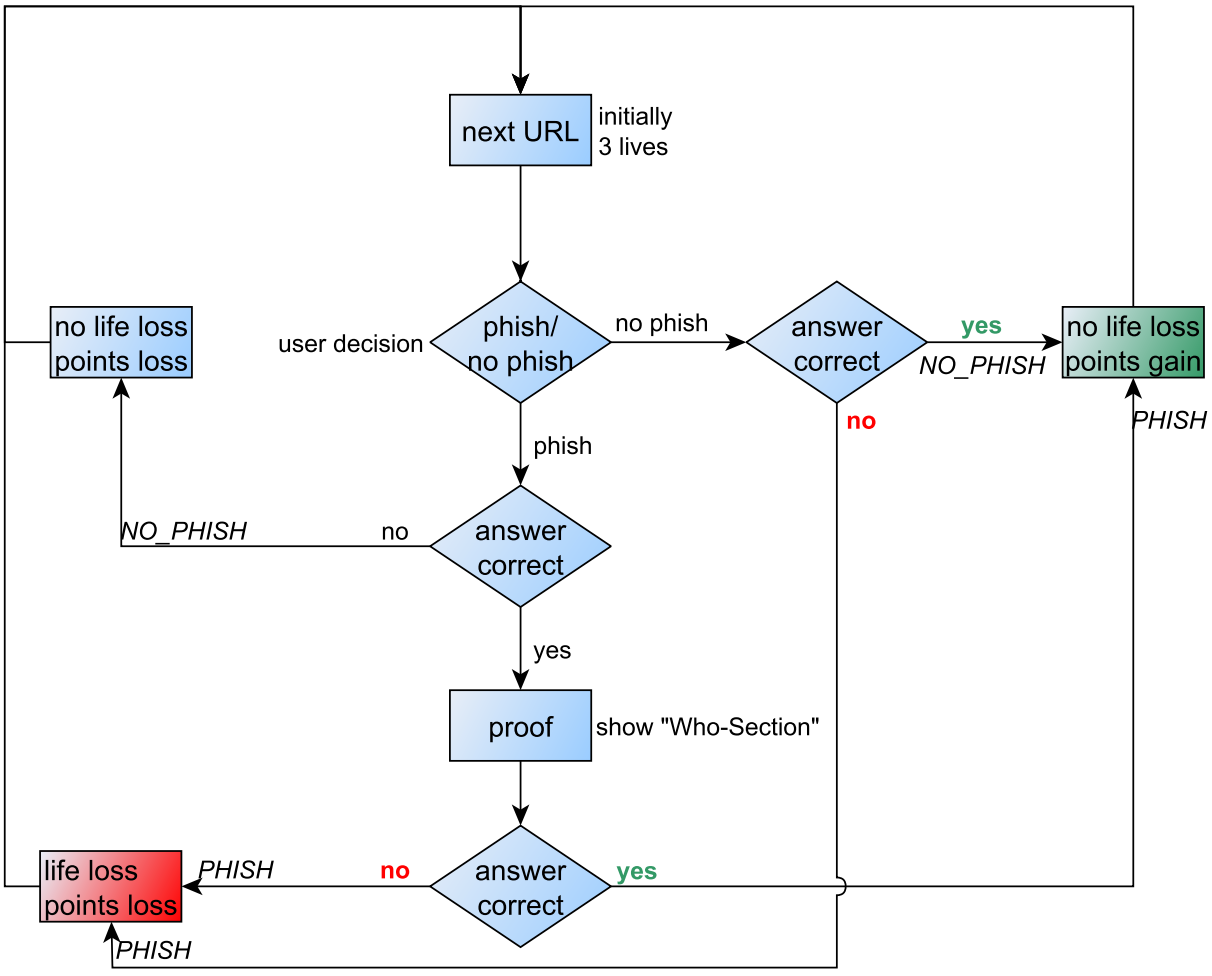

Figure 6: Losing points and lives in the game

8.3 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. 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 the 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 6, 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. This is our final leveling strategy for the app.

8.4 Knowledge Transfer Per Level

This section summarizes the learning objectives of each level. Note that we generally do not use technical term like URL, domain, subdomain, protocol or the like.

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 7 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 7. 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 7.1.1, 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 7.1.1.

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 7.1.1.

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 7.1.1.

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 7.1.1.

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 7.1.1.

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.

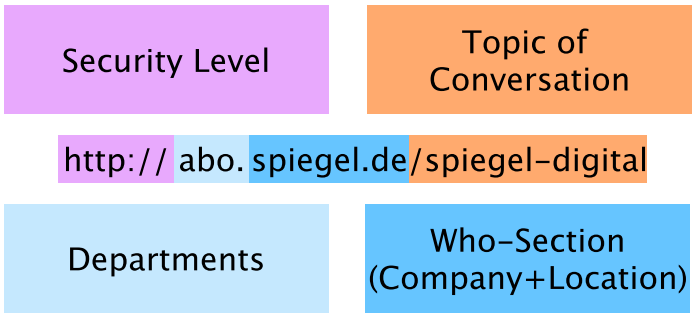

Figure 7: URL components that are communicated to the user

8.5 URL Generation

8.6 Gamification

User motivation

Show Leaderboard Rate

Show Leaderboard Total

...

9 Evaluation

VLL FORMS IN APPENDIX? As a final step the Anti-Phishing Education 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 hypothesis and explain how we are going to measure our statements in order to prove that they are true or false. Finally, we will analyze our results and state our conclusion.

9.1 Study Design

For time reasons and lack of participants we decided to run a "Before and After App" Study with the same groups of people. Specifically, our user study is structured as follows:

1. **General Before-Survey** 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.

2. **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.

3. **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.

4. **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.

5. **General After-Survey** Finally, 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.

9.2 Hypotheses

In order to evaluate the effectiveness and usability of our app we have formulated the following hypotheses for the user study:

1. **Hypothesis 1 - Mistakes** After playing the app, the users make significantly less mistakes in detecting phishing websites compared to before playing the app.

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.

3. **Hypothesis 3 - URL Comprehension** After playing the app the user understands the importance of the second- and top-level domain of a URL as the only criteria to detect phishing websites.

4. **Hypothesis 4 - Good Usability** The app is easy to understand and to use.

9.3 Measurement

In the following we will elaborate on how we are going to measure the statements of our hypothesis and show that they are true or false.

1. **Hypothesis 1 - Mistakes** Correct answers in "Website-Survery After" >> correct answers in "Website-Survery Before"

2. **Hypothesis 2 - URL Based Decision** Number of URL markings in "Website-Survery After" >> number of URL markings in "Website-Survery Before"

3. **Hypothesis 3 - URL Comprehension** Number of marked second- and/or top-level domains of URLs in "Website-Survery After" >> number of marked second- and/or top-level domains of URLs in "Website-Survery Before"

4. **Hypothesis 4 - Good Usability** System Usability Scale (SUS) > 68

9.4 Results and Analysis

9.5 Discussion

9.6 Conclusion

10 Conclusion

This chapter provides a short summary of what we achieved in the scope of this thesis and presents an outlook on future work.

10.1 Conclusion

The objectives of this thesis...

10.2 Findings

10.3 Recommendations

10.4 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 realize due to time and resource limitations.