Iuga *et al. Hum. Cent. Comput. Inf. Sci.* (2016) 6:8
DOI 10.1186/s13673-016-0065-2

Human-centric Computing and Information Sciences

RESEARCH

Open Access

Baiting the hook: factors impacting susceptibility to phishing attacks

Cristian Iuga, Jason R. C. Nurse* and Arnau Erola

*Correspondence: jason.nurse@cs.ox.ac.uk
Department of Computer Science, University of Oxford, Oxford, UK

Abstract

Over the last decade, substantial progress has been made in understanding and mitigating phishing attacks. Nonetheless, the percentage of successful attacks is still on the rise. In this article, we critically investigate why that is the case, and seek to contribute to the field by highlighting key factors that influence individuals' susceptibility to phishing attacks. For our investigation, we conducted a web-based study with 382 participants which focused specifically on identifying factors that help or hinder Internet users in distinguishing phishing pages from legitimate pages. We considered relationships between demographic characteristics of individuals and their ability to correctly detect a phishing attack, as well as time-related factors. Moreover, participants' cursor movement data was gathered and used to provide additional insight. In summary, our results suggest that: gender and the years of PC usage have a statistically significant impact on the detection rate of phishing; pop-up based attacks have a higher rate of success than the other tested strategies; and, the psychological *anchoring effect* can be observed in phishing as well. Given that only 25 % of our participants attained a detection score of over 75 %, we conclude that many people are still at a high risk of falling victim to phishing attacks but, that a careful combination of automated tools, training and more effective awareness campaigns, could significantly help towards preventing such attacks.

Keywords: Phishing attacks, Web security, Human factors, User studies

Background

Phishing in its broadest sense can be defined as a scalable act of deception whereby impersonation is used by an attacker to obtain information from an individual (i.e., the target) [1]. The most common form of online phishing involves the sending of a deceptive message or email which at some point redirects the victims to a legitimate-looking, but malicious website. Once at the website, victims willingly enter their credentials (e.g., usernames and passwords), and maybe even financial information, under the belief that the website they are on is legitimate.

In the last decade there has been significant progress towards preventing and detecting phishing attacks. This progress has come in the form of plug-ins and extensions to detect phishing attempts, machine learning algorithms based on source code features, comprehensive blacklists and whitelists, and others. Moreover, even novel gamification approaches have been proposed such as anti-phishing games, to help individuals learn to identify phishing websites better and faster [2]. After all of these efforts however, researchers still note that there is a lack of existing anti-phishing solutions considered as an optimum [3], while some believe that no static technical defence measure can solely mitigate the threat introduced by user behaviour [4]. The fact is that perpetrators of phishing attacks are continuously adapting their tricks to find novel ways of causing losses to individuals and organisations [5, 6].

From a financial perspective, the impact of phishing attacks is enormous. In the UK alone, the City of London Polices National Fraud Intelligence Bureau (NFIB) and the Get Safe Online security awareness campaign, reported that phishing scams costed victims 174 million during 2015 [7]. In the organisational domain, the most recent FBI Internet Crime Report showed that in 2014 more than $11 million USD was lost due to government impersonation e-mail scams only [8]. Industry can suffer significantly as well, as current studies estimate that the total annual cost of phishing for some average-sized organisations could be around $3.77 million USD [9].

As other research articles have emphasised, if we hope to develop enhanced tools and effective strategies to reduce the successfulness of phishing attacks, we must first know how and why people fall for phishing schemes [10, 11]. This understanding could potentially lead to superior techniques which would help individuals in accurately detecting phishing attacks. The aim of this article, therefore, is to critically investigate the key factors that influence individuals' susceptibility to phishing attacks. We scope our study to explore what happens *after* people are redirected to malicious websites. Do they know how to detect a phishing web page? If yes, why? What factors result in people attributing the phishing label to a certain page?

Our research aims are summarised as follows. Firstly, we will search for correlations between certain characteristics of an individual and how vulnerable they are to falling for phishing attacks once on a malicious website. Secondly, this work will investigate what aspects of a web page are most used for accurately detecting phishing attacks (for example, is the padlock indicator in a browser as crucial as one may expect). Also, we seek to examine what aspects are mostly missed when looking at a malicious website. Finally, we will attempt to correlate the time users spent on each page with the success of detecting phishing attacks, and potentially even draw conclusions on how much time a user would need to be a good manual detector of malicious sites.

Literature review

Though the specific origins of phishing may be deliberated, Symantec notes that the first instances of such attack they witnessed occurred in the 1990s on AOL [12]. The attack vector of choice featured misleading instant messages and emails, which were used to trick users into revealing their AOL passwords, thus granting unauthorised access to their accounts. Since then, phishing has evolved substantially. Initially, there were the generic emails, typically containing several grammatical mistakes, wrongly addressed and directly requesting sensitive information. Now however, with the advent of spear-phishing where attacks target specific people and enterprises, likely for some financial or access gain, attacks have become significantly more tailored [13]. That is, they are professionally composed, directly addressed and cleverly disguise their real intentions. It may be hardly surprising then that studies have found that it takes about 82 seconds for cybercriminals to ensnare the first victim of a phishing campaign [14].

In an attempt to address the broad phishing threat, a variety of countermeasures have been proposed. These seek to focus on the many components of the problem, and advance techniques for mitigation. Mohammad et al. for instance, critically explore the legal proposals and the educational process as a solution [15]. Although they find useful and valuable proposals in this space, they conclude that law enforcement has the downside of a time-inefficient process given the average phishing website's time-to-live, while the educational process is very difficult to implement in practice. On the topic of phishing education and awareness, there are a number of publications. These offer a range of guidelines, training systems, games, and apps [2, 16–18]. Industry has also put forward anti-phishing training suites, one of the most popular being that of Wombat Security [19].

In addition to the education-focused work, many technical prevention and detection mechanisms can found in the literature. Heuristic-based approaches are one of the most novel, and rely on detecting phishing attacks given certain features of emails or web pages. These features are fed into machine learning algorithms which try to correctly identify malicious phishing attempts. A popular browser extension which correlates machine learning with results on search engines is CANTINA+ [20], while an analysis of the performance of different machine learning approaches has also been made [21]. To complement such work, visual aids for phishing prevention have been explored. These target the fact that phishing websites rely heavily on the visual similarity with the actual websites the user intends to visit. The idea behind Security Skins [22] therefore, is that the remote server can identify itself by providing visual aid that is user-specific and easy to recognise. An implementation of this method is Yahoo's Sign-In Seals [23].

Credentials (e.g., passwords) are not the only method for securing online accounts from unauthorised users (such as, after successful phishing attacks). Post-password actions refer to technical methods which are applied after the user enters a correct password. These methods try to detect if anything suspicious is taking place based on aspects such as geo-location and/or the user's behaviour (e.g., two-factor authentication). This type of defence strategy has been referred to by some authors as one of the best client-side defences against hijacking [24]. Finally, there are many comprehensive lists of known phishing sites. Both Google and Microsoft maintain their own blacklists and other known solutions are PhishTank [25] and AIWL [26].

Another area that research has investigated is that of why individuals fall for phishing attacks. Early articles found that some users simply do not look at browser based-cues such as security indicators and the address bar, and even sophisticated users could be fooled by visual deception [10]. Somewhat surprisingly, recent research has found similar issues in the lack of attention paid to security indicators, and poor performance in identifying phishing websites [27]. In terms of user-based factors influencing the likelihood to be successfully phished, studies have discovered that gender, age and exposure to educational materials all have some impact [28]. Other work has explored user characteristics in more detail and identified that user extraversion, trust and submissiveness represent aspects that prevent victims from suspecting phishing attacks [29].

While research works such as those mentioned above have greatly advanced the field of phishing prevention and detection, we believe there are some areas, particularly in the individuals' susceptibility to phishing attacks, still to be addressed. These include, consideration of a broad range of factors (demographic and page-specific), assessing the impact of phishing attacks via browser popups, and more thorough analysis into what areas on web pages users assess to determine page legitimacy. It is these aspects therefore, that form the goals of this paper.

Methods

Overview

To fulfil the aims of this research, we designed and conducted a web-based user study. At its core, the methodology for this study involved: recruiting a set of participants; gathering information (including demographic data) about them; presenting them with sites and requesting that they differentiate between legitimate and phishing pages; and finally, analysing their responses and other related data to draw conclusions. In what follows, we expand on the various components of this methodology.

Using the research aims, we first set about designing the details of the study itself in terms of its flow, the types of sites to be presented and how we would analyse the data. Then we set up the site supporting the study such that when participants accessed the landing page, they were briefed about the scope of the research project and asked to read a short description of the study structure and what was about to happen. If the participants agreed to participate, they were taken to the next page where we gathered basic demographic and user-related information. The questions asked were about participants' age, gender, education level, profession, computer experience, whether they had been targeted or victims of phishing attacks before (and any losses), whether they had received any phishing awareness training, and their perception on how likely they were to detect all phishing attacks in the study.

After submitting demographic information, participants were presented with images specifically designed for this research. These images were harmless representations of phishing and legitimate websites; an important point from an ethical perspective. In order to make the study as close as possible to a real phishing scenario, we implemented a fullscreen page API.[1] Moreover, by hovering their cursor over specific areas of the images, participants were able to identify useful information about the web page including security certificate information, button links, and basic source code. This was setup to mirror real web pages and to allow participants to move their mouse in the areas they were actually inspecting. To capture this mouse movement data and gain further insight into participants' thought process, we used MouseFlow,[2] a website heatmap and session replay tool.

For the actual study, the task of participants was straightforward. To each image they were shown, they were asked, 'Would you enter your login credentials on this webpage?', with two possible answers: 'Yes' or 'No'. Answers were recorded to enable analysis of the data at a later stage. Participants were given no option of going back once they submitted an answer. Lastly, before launching the study, we ran a one-week pilot, with individuals of various backgrounds, and made improvements to the study as necessary.

[1] http://davidwalsh.name/fullscreen

[2] https://mouseflow.com/.

Images design

For our study, we decided to limit the scope, and thus the images and pages used, to a single domain; namely, *facebook.com*. This gave us the opportunity in this particular experiment to focus only on the elements we wanted to change and test against, as opposed to full sites. Facebook was selected as it is the second most visited site in the world according to Alexa,[3] and also, a treasure trove for attackers interested in identity theft, attacking corporations and a host of other threats.

The images displayed to participants alternated between two elements that could be identified by a user in order to detect a phishing attempt: the URL or link and the https indicator (security certificate). Source code fragments were also displayed when hovering the cursor over certain sections in the images but this was only done to reduce the bias towards the elements that needed inspection and provide a uniform experience. Every image was designed to replicate three different browsers (Chrome, Firefox, Internet Explorer) to offer participants a familiar interface where they could focus on detecting phishing. As a result, we designed a total of 39 images, 13 images for each browser, and organised them into three different groupings (hereafter, contexts) of pages each.

Context 1 displayed six images based on the login page of Facebook. Only the `https` and `URL` aspects were modified, and for ease of reference, we summarise those modifications in Table 1.

In detail, the first image in this context had a URL which denoted a regular, unprotected channel of communication but, more importantly, it had three 'o' letters instead of two. As such, we considered this page to be a malicious one. Figure 1 displays this page and serves as a visual example of pages from this context. The second image still presented an unprotected channel of communication, but had the correct spelling of Facebook. We considered this page to be legitimate since `http only` does not necessarily imply phishing. In fact, at the moment, not many pages offer SSL/TLS certificates with only 23.4 % of the 150,000 most popular websites being reported by the trustworthy internet movement to use SSL/TLS [30]. Above all, this image could lead to interesting findings about how people perceive websites, and particularly phishing sites as it relates to the use of security certificates.

The third image was of a correct Facebook login page over `https` and had a redirect variable at the end. Its purpose was to test whether participants would know what to make of the content of the URL that followed the correct domain. This page was regarded as legitimate. The fourth image was designed to be a somewhat more challenging one. We hypothesised that in the first image participants may focus on the middle section of the link (3 o's), in the second they may focus on the beginning part (`http only`), and in the third one, at the end of the link (`?_rdr`). For this image therefore, we corrected only the beginning and end, but reintroduced the triple o's; thus making it a phishing page.

The fifth image was of the standard correct `https` Facebook login page, while the sixth image introduced a facebook subdomain on *secure.com* with a longer link. This latter image was considered to be a phishing attempt, and our goal was to investigate

[3] http://www.alexa.com/topsites.

Table 1 Images of web pages—Context 1

#	Detailed characteristics	URL
1	http only + faceboook (3 o's)	http://www.faceboook.com
2	http only	http://www.facebook.com
3	https + redirect (rdr)	https://www.facebook.com/?_rdr
4	https + faceboook (3 o's)	https://www.faceboook.com
5	https	https://www.facebook.com
6	https + subdomain secure.com + long link	https://facebook.secure.com/?_apiKey=[...]

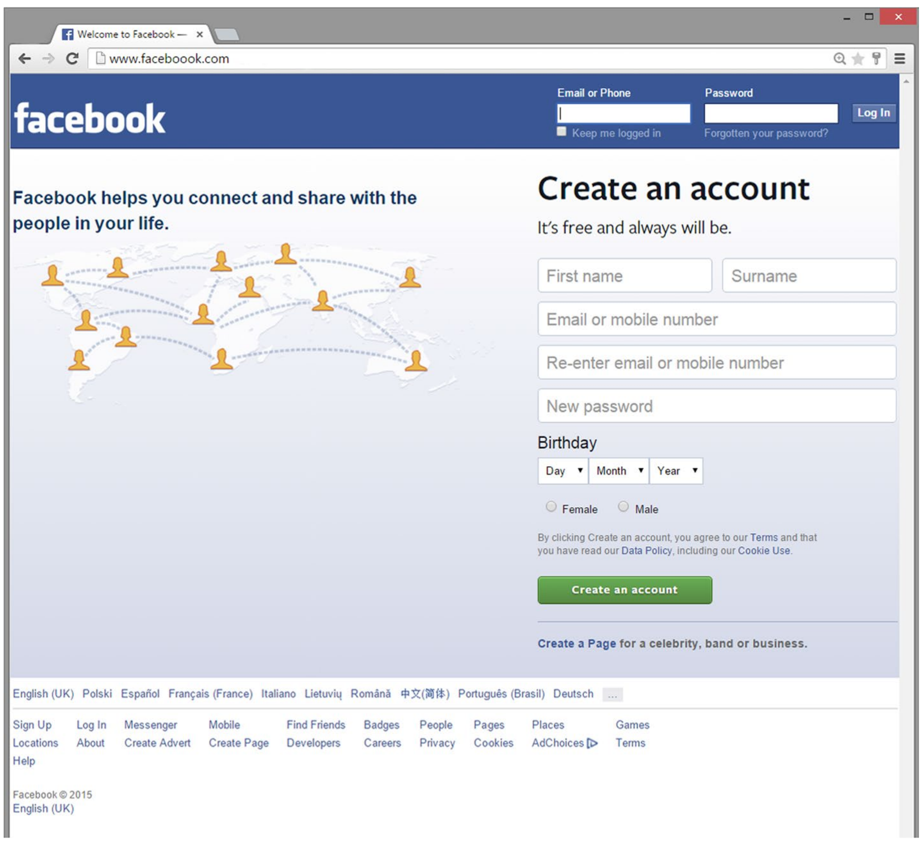

Fig. 1 First web page from Context 1. This screenshot represents Page #1 in the study, and is defined by the following characteristics: http only and faceboook (3 o's)

whether individuals would know the difference between the domain and subdomain of a given web page.

Context 2 displayed three images based on a Facebook page that warned participants that their passwords needed to be changed. An example of such a page can be seen in Fig. 2, and a summary of the characteristics of the pages is available in Table 2. Before actually presenting this new set of images, we notified participants via a short text about the change of context (i.e., types of screens). The text asked participants to now assume that they followed a link in a message purporting to be from Facebook. The

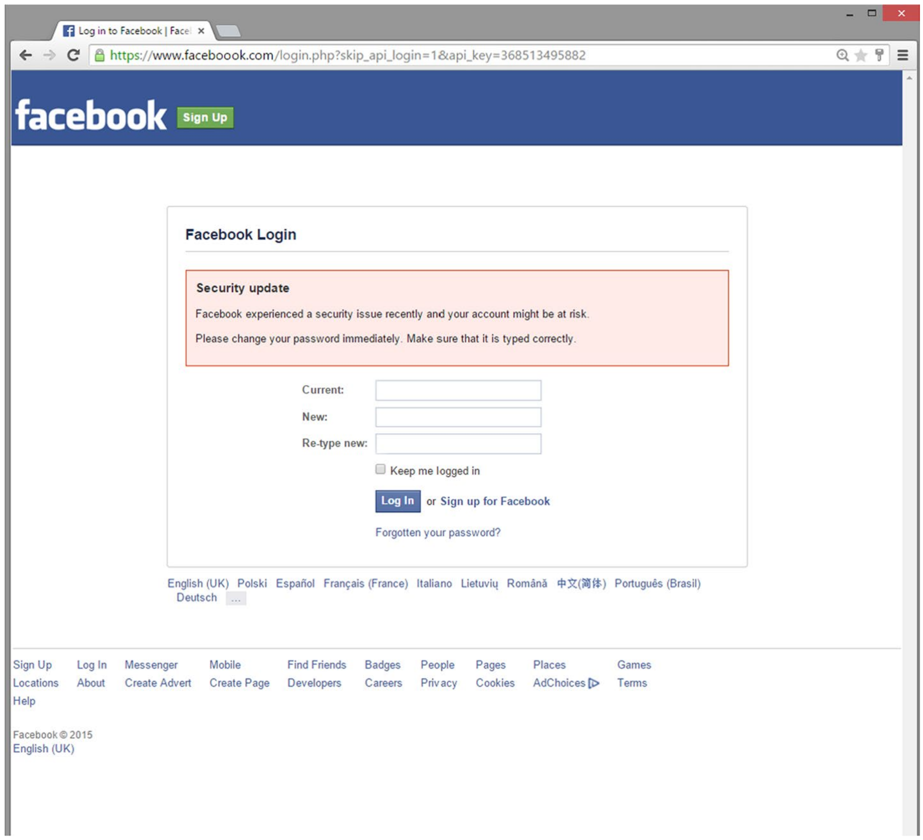

Fig. 2 First web page from Context 2. This screenshot represents Page #7 in the study, and is defined by the following characteristics: faceboook (3 o's) + long link

Table 2 Images of web pages—Context 2

#	Detailed characteristics (https + update alert)	URL
7	faceboook (3 o's) + long link	https://www.faceboook.com/login.php?[...]
8	IP + long link	https://128.237.239.92/facebook/login.php?[...]
9	long link	https://www.facebook.com/login.php?[...]

message informed them of a security breach and thus, the need for them to login to their accounts and change their passwords.

The first page in this context again used the misspelled *faceboook.com* and had a long URL/link as well; this was considered a malicious site. The next image was of a phishing page with an inaccurate internet protocol (IP) URL and a long link, while the third image of this context represented a legitimate, secure Facebook login which also had a longer link. Our objective with these images was to assess whether the security notification might have a some notable impact on participants' behaviour.

Context 3 presented the most varied layout thus far, displaying four images with pop-ups shown on two web domains which required authentication in order to access the full content. The two domains used were *wsj.com* (i.e., the Wall Street Journal) and *catvideooftheweek.com* (i.e., a miscellaneous cat video website). For a summary of the characteristics of the pages in this context, see Table 3. In order to keep participants engaged, we introduced the context with a short piece of text. On this occasion, we explained that some websites require user login to access their information, and for this context participants should assume that they were somewhat interested in such content. They were then told that the question of entering their credentials information would now refer to the pop-up box.

The first pop-up introduced a *facebook* subdomain of the Wall Street Journal domain, *wsj.com*; therefore representing a phishing site. A key goal in this context was to assess whether there was any difference between individuals' perception of a regular subdomain link versus the pop-up subdomain link. The second image of this context displayed a correct Facebook pop-up on *wsj.com*, while the final two images depicted phishing attempt pop-ups with the same domain as the parent page: *catvideooftheweek.com*. Figure 3 shows an example of this context's images, with Page 12. The only difference between the last two phishing attempts was that one of them had an unprotected communication channel while the other offered a secure link for the pop-up.

Summarizing, each participant was shown 13 images, of which 7 were considered phishing attempts while the others were representations of legitimate websites. We paid special attention to make the images as similar as possible to a real desktop, avoiding any bias that our web page could introduce.

Ethics, recruitment and data collection

The study was launched on May 2015 and over a period of a month we gathered participants' information. Our recruitment drive mainly involved advertising the study through social media and email, using research and industrial contacts. In the study itself, we collected the following data: participant demographics and user-related data; responses to the questions regarding the presented images; time spent on each page; the cursor movement on these pages; and optionally, email addresses if participants wished to receive the study's results and further details about phishing prevention. As mentioned, the cursor tracking information was gathered and stored by a third party tool (MouseFlow) and allowed us to collect more quantitative information related to participants page actions. This study received ethical clearance from the University of Oxford Research Ethics Committee to be conducted.

Table 3 Images of web pages—Context 3

#	Detailed characteristics (popup)	URL
10	https + subdomain wsj.com + long link	https://facebook.wsj.com/login.php?[...]
11	https + facebook + long link	https://www.facebook.com/login.php?[...]
12	cats + http only same domain + long link	http://catvideooftheweek.com/facebookLogin.php?[...]
13	https + cats + same domain + long link	https://catvideooftheweek.com/facebookLogin.php?[...]

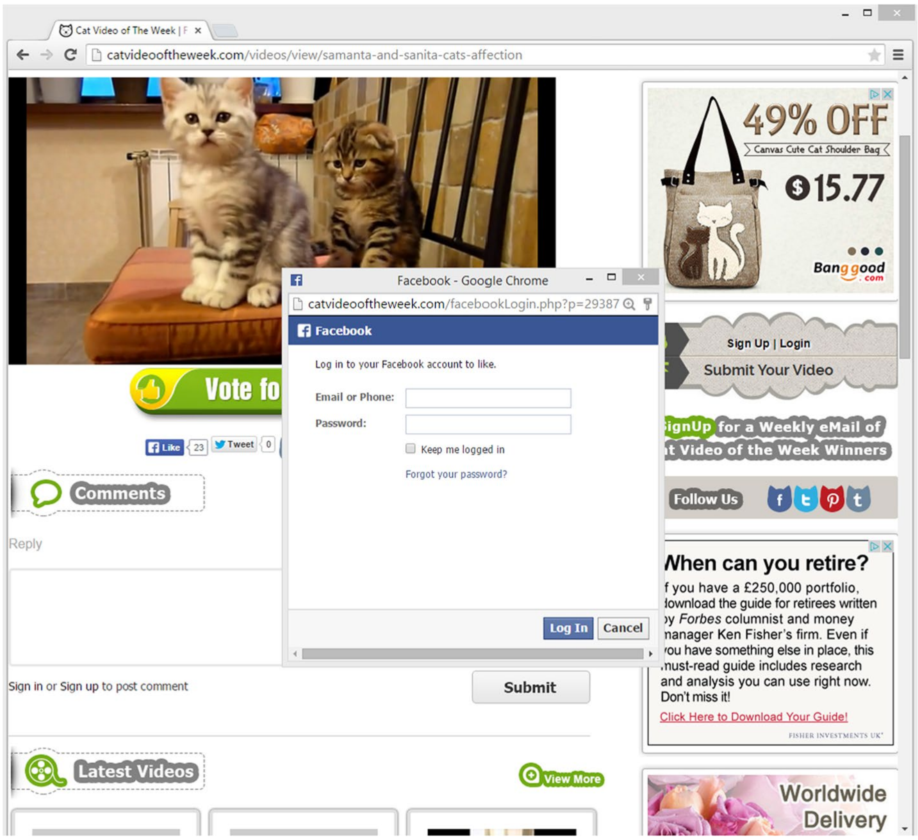

Fig. 3 Third web page from Context 3. This screenshot represents Page #12 in the study, and is defined by the following characteristics: cats + http only same domain + long link

Data analysis

Analysing the data gathered involved several tasks. Firstly, we conducted an overview analysis which was coded using PHP, MySQL and a charting API called Highcharts.[4] We gained an overview of our population sample, of the time spent in our study, and also of other individual page results such as the rate of success and failure for each page (i.e., detection score). Secondly, using IBM's SPSS software,[5] we applied an ANOVA (analysis of variance) test for every user feature to determine whether any statistically significant difference of means among groups existed; here, groups relate to categories of the independent variables, such as male and female. We then analysed the probability (p-value) of obtaining the data assuming the null hypothesis which states that all population means are equal. A p value <0.05 was considered as significant. Moreover, we ran a Tukey's post hoc procedure to compare all different combinations of the treatment groups.

When reporting statistical results in the next section, we present details of the F-ratio and the degrees of freedom from which it was calculated, the effect size η^2, and the p value. Finally, since cursor tracking information was collected by MouseFlow, we

[4] http://www.highcharts.com/.

[5] http://www-01.ibm.com/software/analytics/spss/.

inspected the heatmaps generated by this tool. This was done for every individual page to attain an understanding of the summarised cursor movements on these pages. Analysis and conclusions pertaining to this data were based on our visual interface.

Results and discussion

Participant sample

In one month after the launch of our study, 382 individuals fully completed the study, while 556 individuals had passed the initial demographic questions page. For our analyses, we only considered those who had fully completed the study. An overview of the demographics and information of the participant sample is next.

In total, we registered 240 male participants (62.83 %) and 142 female participants (37.17 %). The age group 18–24 accounted for 50 % of the sample but we also had five participants over 60 years old. Most of the participants were either currently studying for a university degree or had already completed one, with only 5.76 % selecting high school or less. In terms of location, out of the 21 countries we had participants from, Romania and the UK hold more than 80 % of the total, with 61.78 and 21.20 % respectively; these were the two countries where the recruitment drive was the most significant. 53.14 % of our participants were employed and 43.19 % were students. Most of our users had more than 6 years of PC usage. Overall, 65.71 % of our participants had been subjected to phishing attacks with under 6 % of these declaring financial or credentials losses. 19.90 % do not know if they were exposed to phishing attacks, while 14.40 % think they were never subjected to such attacks. Only 21 people across the sample had phishing training before, which accounts for 5.50 % out of participants.

Plenty of phish in the sea

The first main finding from our study was that phishing still is a significant concern. We found that the average success rate of detecting a phishing site in our study is 65.63 % with only six participants attaining a 100 % score. Even more concerning, only a total of 26 people scored more than 90 % correct. This situation acts to confirm prior research and reiterate the seriousness of the problem. One question that occurred to us having seen such a score, was whether our decision to consider the Page 2 in the study as legitimate dramatically affected the results. Indeed, we did mention in our image design section that this page would be one of particular interest. As such, we re-ran the analysis considering the `http only` Facebook login image as a phishing page. The results showed that the average success rate was slightly lower in this configuration (63.83 %). However, an improved result was that nine participants now scored 100 % correct.

We also analysed the data of participants who performed poorly in the study. If a participant would have responded to all answers with a 'no', their detection score would have been 53 %. However, 37 of our participants received an even lower score than this which denotes a worrying reality. As we have witnessed, such attacks can cost corporations significant amounts [9]. Our results seem to agree with a recent study which showed that people fall for the most obvious phishing pages 3 % of the time, for the average phishing pages 14 % of the time, while the most believable phishing pages trick users 45 % of the time [24].

Confidence level

Before presenting the images of this study to our participants, we asked them how likely it was that they would detect all the phishing pages in the study. They had five options to choose from: *very low probability, low probability, sort of random, high probability, and very high probability*. On completing the study, we asked the same question again, therefore recording the confidence level before and after the study. After analysing the data, we discovered that only 11.52 % of our participants reported a higher confidence level after completing the study, while 39.01 % reported a decreased confidence level. The remaining 49.48 % selected the same value of confidence level with almost all of them reporting a level of *sort of random* or below. We hypothesise that the high percentage of decreased confidence levels along with the high percentage of participants that entered their email addresses to find out more about phishing, could be an indicator that the study successfully challenged participants to take phishing more seriously.

Statistically, ANOVA results suggest that there exist correlations between the initial confidence level and the detection scores, $F(4, 377) = 12.664$, $\eta^2 = 0.11$, $p = 0.000$, as well as between the final confidence level and the detection scores, $F(4, 377) = 13.651$, $\eta^2 = 0.12$, $p = 0.000$. From a comparison of the mean scores of the final confidence level in particular, we noticed that participants were much better at judging their performance.

Demographic feature analysis

As mentioned earlier, we collected user demographics and other information across several features. For each of these features, we ran one-way ANOVA tests in order to determine if they were linked with the detection scores or with the time each participant took in the study. Some of the main results of the ANOVA are listed in Table 4. Here, due to space restrictions, we only provide details about the features that were statistically significant for either the detection score or the time spent.

Similar to other studies, our results show that there is a statistically significant difference between males and females in detecting phishing web pages, $F(1, 380) = 13.467$, $\eta^2 = 0.034$, $p = 0.000$. Sheng et al. relate this difference to the supposed less technical experience that women have compared to men [28]. PC usage also has an impact on phishing detection. In our analysis, this feature was statistically significant for both the detection score, $F(3, 378) = 5.495$, $\eta^2 = 0.041$, $p = 0.001$, as well as the time spent on the study, $F(3, 378) = 2.837$, $\eta^2 = 0.022$, $p = 0.038$. This suggests that a longer interaction with computers (and probably the World Wide Web)

Table 4 Demographic features analysis

Demographic feature	Detection score correlation	Time correlation
Age	No	Yes
Gender	Yes	No
Education	No	Yes
Country	Yes	No
Profession	No	Yes
PC usage	Yes	Yes
Browser	No	No

contributes towards a better understanding of phishing attacks and detection strategies, therefore helping users to identify phishing web pages faster and more accurately.

The time spent on all the pages in our study is significantly correlated with age, $F(8,373) = 4.356$, $\eta^2 = 0.085$, $p = 0.000$ and profession, $F(3,378) = 4.682$, $\eta^2 = 0.035$, $p = 0.03$. One reason for this could be the possibility that older participants tend to be less impulsive and prefer to properly assess a site before making their decision. A similar reality may be true for individuals who are employed (as opposed to students, for instance); that is, general awareness or workplace training could have influenced them to think more carefully about such actions. This also related to another point: unemployed participants spent on average almost the same time as students which account for the lowest amount of the time taken to complete the study. While we were unable to find a specific reason for this, we posited that lower risk awareness or potentially even education may have been a factor. Lastly, although in our analysis the country feature is statistically correlated with the detection scores, the population sizes of most countries are very small compared to those of the United Kingdom and Romania. As such, we are unable to definitively conclude that the country impacts detection scores.

Critical pages

Having discussed general study findings, we now focus our attention to assess the critical web pages. We define a critical web page as the page on which the set of participants failed more than they succeeded. From the thirteen images shown, Page 1 (P1), Page 4 (P4) and Page 11 (P11) received more incorrect answers than correct ones. Figure 4 presents an overview of the numbers of successes and failures for each page. The green colour represents the number of correct answers given for each page while the red colour accounts for the incorrect ones.

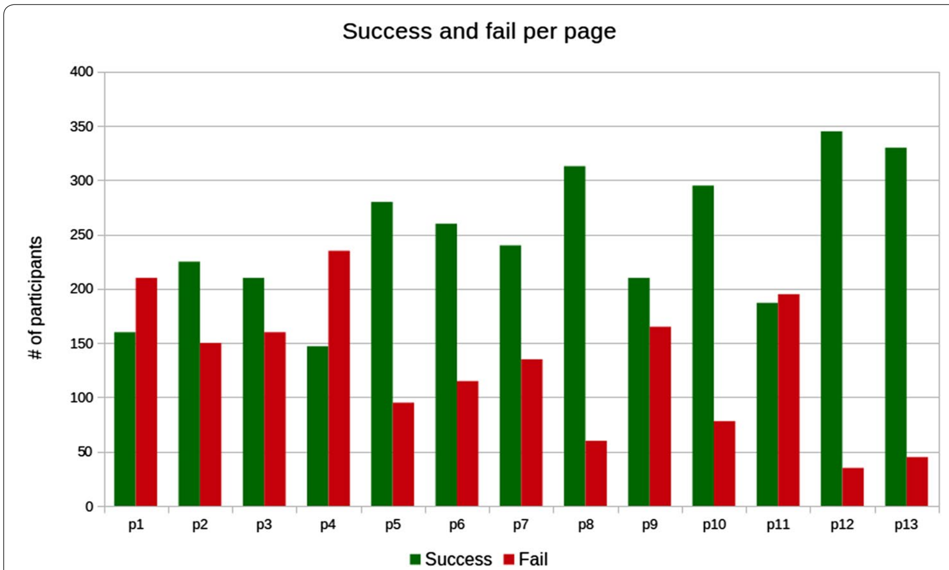

Fig. 4 Successes and failures per page

Despite the fact that the first page in our study had the highest average time spent on a page (25 s), 214 participants (56 %) entered an incorrect answer on it. This seems to suggest that although participants allocated what seemed to be enough time to make a decision, they were still unable to correctly identify it as a phishing web page. Since this page was `http only` and displayed a misspelled "faceboook" (3 o's), one question we asked is whether participants actually inspected the URL before giving their answer.

Therefore, we examined the cursor movement heatmap presented in Fig. 5. We can clearly see in the heatmap that most of the participants placed the cursor over the triple o's. Also, the document icon which is specific for `http only` web pages accounted for even more of inspection. It might therefore be hypothesised that the reason for so many failures on this web page does not relate to participants not inspecting the URL or the security indicators displayed in the browser. Whilst we were unable to ascertain a reason for this, we did wonder whether the answer was in a lack of knowledge about what `https` meant or the participant's attention to detail. Generally speaking however, this is quite concerning, as participants are seemingly overlooking key indicators that may be used to detect phishing attacks.

Although our participants clearly considered Page 4 to be easier and spent an average of only 10 seconds on it, we designed it as quite the contrary. This page tested the hypothesis that the anchoring effect applies also to phishing. By definition, the anchoring effect is a cognitive bias that influences someone to rely too heavily on the first piece of information they receive. As such, we tested whether the first three pages would influence the zones that participants inspected on the fourth one. Regarding the analysis of the first three pages, in the first two images users focused on the beginning part (`http only`) and the middle section of the link (3 o's), while the third page obtained more focus on the end of the link (`?_rdr`). However, in the fourth image we corrected only the beginning and the end of the link but reintroduced the triple o's in the middle.

Looking then at the cursor movement heatmap of the fourth page (Fig. 6), we can easily see that the middle part of the link was not inspected as much as it should have,

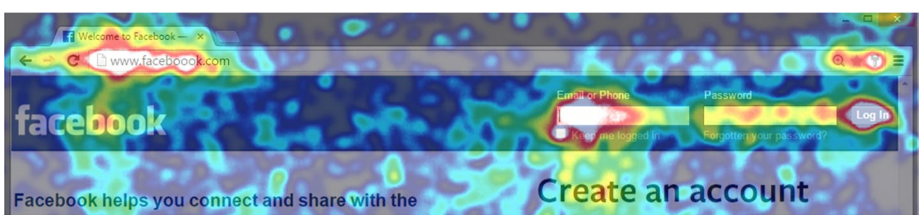

Fig. 5 Heatmap for Page 1 (cropped)

Fig. 6 Heatmap for Page 4 (cropped)

therefore somewhat supporting our hypothesis. As a likely result of the cognitive bias, this page accounts for the highest failure rate of all the pages in our study. 61 % of our participants submitted an incorrect answer and all of them spent an average of 10 seconds to inspect the image.

The last critical page, Page 11, represented a page where a legitimate Facebook pop-up is displayed and logging in is required to show the full content of an article. Here, 51 % of our participants incorrectly labelled it as a phishing web page with an average time of 12 seconds spent to investigate its correctness. Whilst on the positive side, participants appear to have chosen the safe option (potentially when they are not able to confidently determine the legitimacy of a page), this finding does raise questions about how individuals perceive pop-ups. Pop-ups are a valid way of web developers using federated identities to allow sign-in across websites, but being visually presented 'over' other sites and displaying slightly different, minimalistic interfaces, may act to confuse users online.

Where the mighty fall

Another question that we asked during the individual page analysis is this: where do participants who only miss one page fail? Perhaps this will show what factors are mostly missed by people even when they have good knowledge of cues for detecting phishing attacks. Similar to before, whilst we maintain that Page 2 is legitimate, we also wanted to consider the scenario if it was regarded as a phishing site and the impact on the study results, if any. Figure 7 shows the number of participants who missed only one page when considering Page 2 to be legitimate in yellow, and the number of participants who missed only one page when considering Page 2 as a phishing attack in blue. In total, we have 20 participants in yellow and 13 in blue.

The first aspect to notice is that no matter the context, pages 2, 9, 10 and 11 account for wrong answers. Page 2, in particular, proved rather challenging as participants struggled with whether or not `http` alone represented a phishing attack. In both situations, the

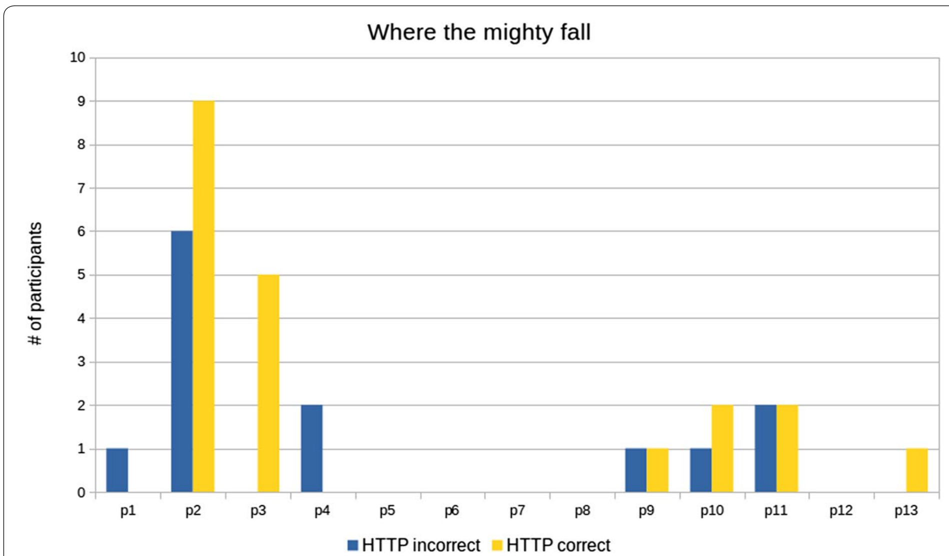

Fig. 7 Number of fails on each page for the best participants

number of fails are the highest over all the pages, being 6 and 9 respectively. This highlights an interesting question regarding the perception of `https` and what it really means. The context of Page 9 is a legitimate Facebook page which asks users to update their passwords. Page 10 (phishing) and Page 11 (legitimate) represent Facebook pop-ups which supposedly after a successful login would allow access to content on Wall Street Journal (*wsj.com*).

To add to this analysis, Fig. 8 presents the average time spent on each page by the most successful participants. When assessing these data, we can see that the failures on Pages 2, 9, 10 and 11 can not necessarily always be explained by less time spent to inspect these pages. Therefore, we believe this finding may highlight at least two important points. Firstly, experienced users do not necessarily trust correct security indicators or notifications, and secondly, phishing attacks which rely on malicious pop-ups may be harder for users to detect.

Unfortunately, the first of these findings suggests that the slew of phishing attacks over the last decade may have impacted the trust that some users are placing in security cues. This could have a further impact on companies themselves in the long term, if users view online services as untrustworthy. Moreover, results suggest that when using pop-ups, perpetrators of phishing attacks could be more successful in their objective of deceiving users. On the positive side, however, if people knew how to detect malicious links in general, they could apply the same information for preventing phishing attacks which rely on pop-ups. Therefore, the number of cues that one would have to learn in order to correctly classify web pages does not increase when considering this new channel of phishing attacks. Overall however, it was encouraging to see that in two of the three cases, participants failed because they were more cautious than not.

Anti-phishing training impact

As mentioned before, 21 participants received anti-phishing training at some point prior to taking this study. Our results show that the average detection score for the

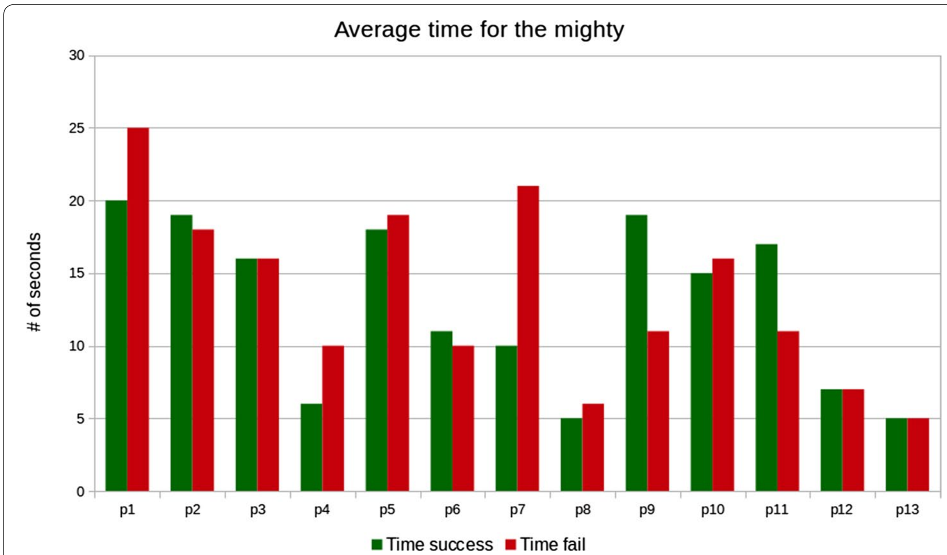

Fig. 8 Average time analysis for the best participants

participants who received anti-phishing training is 67.10 %, while the other category had a detection score average of 65.55 %. Also, none of those who had received phishing training before our study managed to score 100 %. We suspected that this could be related to Page 2 being considered as a correct one but when we changed it to a phishing page, the maximum detection score for individuals who received training decreased even further, from 92 to 85 %. Lastly, we found that participants, who had phishing training before, took on average almost 20 s more to complete the tasks in the study.

While it is tempting to use these findings to posit that anti-phishing training is not effective, that would not be prudent given the small sample size of those who had such training; this small size is also our reason for not conducting ANOVA tests. Unfortunately, we did not collect additional information on how recently participants had their training or the details of such training (e.g., a phishing game, printed materials, class session), which also prevents any high-level insight on their perspective. This is an area for us to expand on in future work as studies have shown that training can have an impact on performance [17].

Phishing loss impact

Although most of our participants reported that there was no loss after being subjected to phishing attacks, 4.19 % reported that they lost their credentials after such attacks and 1.31 % reported financial losses after being phished. We had hypothesised that participants who lost personal information during phishing attacks would be better at detecting such attacks than the group of participants who have not because of their potential increased awareness. However, this was not the case after simple comparison across groups. We do caveat this finding with the fact that the group size was quite small, and therefore, findings are hardly conclusive.

We also attempted to gather some understanding as to whether there was any difference in detecting phishing between individuals who lost only credentials and those who lost financial data during phishing attacks; potentially the latter category would be much more prepared since the attacks reached their finances. Nonetheless, although participants who had financial losses had a better average than those who lost only credentials during phishing attacks (a 10 % difference), small sample sizes prohibit us from conducting proper tests for significance. In the future, we may look to explore this aspect further by specifically attempting to recruit a larger set of individuals who have suffered loss due to a phishing attack.

Limitations

While this research study has achieved its aims and identified key factors that influence individuals' susceptibility to phishing attacks, there are two main limitations to be noted. The first limitation is that there was no smartphone support for the study. Although we ran a pilot study beforehand, our participants were all manually selected and used a PC to access and complete it. Immediately after the first post on Facebook to advertise the study, we noticed that many people accessed our landing page through mobile devices but unfortunately our study was not suited to be taken on a mobile device.

The second limitation was related to the use of mouse tracking versus eye tracking. Chen et al. present a study on the relationship between gaze position and cursor position on a computer screen during web browsing [31]. They concluded that mouse tracking is an inexpensive and scalable technique that has merit as an alternative to eye-tracking systems, especially in usability evaluations on the web. In order to have even more meaningful mouse tracking data, we gave the users an incentive to move the mouse in the areas they were inspecting by showing pop-up boxes with additional information (thereby mirroring typical browser behaviour). However, we do acknowledge that mouse tracking may not be as exact as eye tracking. For further work therefore, we could consider conducting the same study on a smaller sample using eye tracking, to examine the differences if any.

Conclusions and future work

Although in the last decade significant progress has been made to prevent phishing attacks, it seems that by continuously adapting their bait, attackers are still able to trick and exploit some users. This study shows that when on a malicious web page, people will fall for phishing schemes with an average success rate of detecting phishing of 65.63 %. Moreover, in our study, only six participants attained a 100 % detection score which accounts for 1 % of the total population sample. By running statistical tests on our data, we identified that gender and the years of PC usage have a significant impact on the detection scores. This acts to reinforce some aspects of earlier literature.

While conducting an individual page analysis which correlated detection scores, time spent on each page, and cursor position heatmaps, we reported that many of the participants who failed to correctly identify only one page in our study had missed phishing attacks which relied on pop-ups. Therefore, we concluded that these kinds of attacks are harder to detect; future analysis may highlight that the reputation of the parent page of the pop-up might have an impact on correctly spotting phishing attacks. Also, the study showed that the anchoring effect is present when detecting phishing. Heatmaps supported our hypothesis, thus suggesting that recent history has an impact on the ability of people to detect phishing web pages.

To briefly reflect on our research in the context of existing literature, we believe that we have contributed in two important ways. First, we extend current knowledge on factors that influence phishing susceptibility (e.g., gender, PC usage, user extraversion, trust and submissiveness [28, 29]) by identifying new aspects such as the use of pop-up based attacks and the psychological anchoring effect. This is extremely useful as it highlights new areas where individuals may be vulnerable and therefore, need more protection. We also highlight that there are numerous factors that we demonstrated do not impact an individual's likelihood to be phished; this can inform and allow better scoping of future research. Next, our research further supported current literature, and in many ways, actually provides clear evidence to explain previous findings. For instance, literature has found that users simply do not look at browser based-cues such as security indicators and the address bar, hence fall for attacks [10]. We provide clear evidence in support of this finding, and the fact that some users may not even understand the security implications of the cues that they are looking at; this was achieved through the tracking and analysis of exactly where the user moves their mouse as they consider phishing and legitimate web pages.

There are many avenues for future research in this space. One immediate area to extend this current study could be exploring the susceptibility to phishing attacks in mobile device interfaces. Studies have already began to consider the general problem of phishing on mobiles and the increasing number of possible attacks [32]. However, there is need for more work on understanding the key factors that impact users' susceptibility here. For instance, we could investigate whether the same factors arise, or different ones exist; one main point of course, being the size of such devices. This could be interesting especially if we consider the Internet-of-Things and the various sorts of low resolution devices, and phishing attacks that might be launched. Then, the goal would be to create enhanced approaches to respond to such attacks, whether they be technical or educational-based.

To consider on our work more broadly, the aim of this research was to understand the susceptibility of individuals to phishing attacks and the key related factors, to then use them for better prevention and detection. From what we have learned in this article, there are two main areas that we will aim to focus on in the future. Firstly, we will seek to design and develop technical solutions by way of web browser extensions. The sensitivity of these extensions for providing warnings of potential phishing sites will be adapted based on user's characteristics; i.e., the extensions would be such that the more vulnerable types of users potentially receive more, and increasingly tailored warnings. We will take inspiration from research on discovering likely phishing attacks [3], and developments on effectively communicating warnings to users [33, 34].

Secondly, we aim to produce highly focused educational and training approaches that consolidate the various contributions thus far on the most effective ways to assist users in avoiding phishing attacks. Even though our study was unable to clearly identify phishing training as useful, largely due to the small sample size, other research on the effectiveness of training has been quite promising [17, 35, 36]. In both these areas, we aim to draw on the wide array of research on users in terms of risk understanding and communication, and mental models of security [37, 38]. Whilst we do not believe phishing is a problem that can be solved easily, there are several steps that can and must be taken in an attempt to address it before the range and impact of attacks become even more significant.

Authors' contributions

This article is the product of dissertation research at the University of Oxford by CI, supervised by JN and AE. As a result, a majority of the core research and experimentation was conducted by CI, with JN and AE providing direction and guidance during the research. All parties assisted significantly in the journal manuscript drafting stage. All authors read and approved the final manuscript.

Authors' information

Cristian Iuga holds a MSc in Computer Science from the University of Oxford, and is currently a passionate Technology Entrepreneur in Romania. Jason R. C. Nurse is a Postdoctoral Cyber Security Researcher in the Department of Computer Science at the University of Oxford, and Junior Research Fellow in Wolfson College, Oxford. Arnau Erola is a Postdoctoral Cyber Security Researcher in the Department of Computer Science at the University of Oxford.

Competing interests

The authors declare that they have no competing interests.

Received: 27 January 2016 Accepted: 3 May 2016
Published online: 10 June 2016

References

1. Lastdrager EE (2014) Achieving a consensual definition of phishing based on a systematic review of the literature. Crim Sci 3(1):1–10. doi:10.1186/s40163-014-0009-y
2. Sheng S, Magnien B, Kumaraguru P, Acquisti A, Cranor LF, Hong J, Nunge E (2007) 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. ACM, New York, pp 88–99. doi:10.1145/1280680.1280692
3. Zeydan HZ, Selamat A, Salleh M (2014) Survey of anti-phishing tools with detection capabilities. In: International symposium on biometrics and security technologies (ISBAST), 2014. IEEE, New York, pp 214–219. doi:10.1109/ISBAST.2014.7013124
4. Dodge RC, Carver C, Ferguson AJ (2007) Phishing for user security awareness. Comput Secur 26(1):73–80. doi:10.1016/j.cose.2006.10.009
5. Furnell S. Still on the hook: the persistent problem of phishing. Comput Fraud Secur (10). doi:10.1016/S1361-3723(13)70092-7
6. Vishwanath A, Herath T, Chen R, Wang J, Rao HR (2011) Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. Decis Support Syst 3:576–586. doi:10.1016/j.dss.2011.03.002
7. Raywood D (2016) Phishing costs UK £174 Million in 2015. http://www.infosecurity-magazine.com/news/phishing-costs-uk-174-million-in/. Accessed 4 Jan 2016
8. FBI (2014) 2014 Internet crime report. https://www.fbi.gov/news/news_blog/2014-ic3-annual-report. Accessed 4 Jan 2016
9. Ponemon Institute (2015) The cost of phishing and value of employee training. http://www.rsaconference.com/blogs/how-much-will-that-phishing-trip-cost-you. Accessed 4 Jan 2016
10. Dhamija R, Tygar JD, Hearst M (2006) Why phishing works. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, pp 581–590. doi:10.1145/1124772.1124861
11. Downs JS, Holbrook MB, Cranor LF (2006) Decision strategies and susceptibility to phishing. In: Proceedings of the second symposium on usable privacy and security. ACM, New York, pp 79–90. doi:10.1145/1143120.1143131
12. Symantec Corporation: A brief history of phishing (2007). http://www.symantec.com/connect/blogs/brief-history-phishing-part-i. Accessed 4 Jan 2016
13. Nurse JRC (2015) Exploring the risks to identity security and privacy in cyberspace. XRDS Crossroads 21(3):42–47. doi:10.1145/2730912
14. BBC (2007) Phishing catches victims 'in minutes'. http://www.bbc.co.uk/news/technology-32285433. Accessed 4 Jan 2016
15. Mohammad RM, Thabtah F, McCluskey L (2015) Tutorial and critical analysis of phishing websites methods. Comput Sci Rev 17:1–24. doi:10.1016/j.cosrev.2015.04.001
16. Jansson K, von Solms R (2013) Phishing for phishing awareness. Behav Inform Technol 32(6):584–593. doi:10.1080/0144929X.2011.632650
17. Canova G, Volkamer M, Bergmann C, Reinheimer B (2015) Nophish app evaluation: lab and retention study. In: NDSS workshop on usable security
18. Hale ML, Gamble RF, Gamble P (2015) CyberPhishing: a game-based platform for phishing awareness testing. In: Proceedings of the 48th Hawaii international conference on system sciences (HICSS). IEEE, New York, pp. 5260–5269. doi:10.1109/HICSS.2015.670
19. Wombat Security Technologies (2016) Anti-Phishing training suite. https://www.wombatsecurity.com/suggested-programs/anti-phishing. Accessed 4 Jan 2016
20. Xiang G, Hong J, Rose CP, Cranor L (2011) Cantina+: a feature-rich machine learning framework for detecting phishing web sites. ACM Trans Inform Syst Secur 14(2):21
21. Miyamoto D, Hazeyama H, Kadobayashi Y (2009) An evaluation of machine learning-based methods for detection of phishing sites. In: Advances in neuro-information processing. Springer, Berlin, pp 539–546. doi:10.1007/978-3-642-02490-0_66
22. Dhamija R, Tygar JD (2005) The battle against phishing: Dynamic security skins. In: Proceedings of the 2005 symposium on usable privacy and security, ACM, New York, pp 77–88. doi:10.1145/1073001.1073009
23. Yahoo! Inc (2016) Yahoo personalized Sign-In seal. https://protect.login.yahoo.com/. Accessed 4 Jan 2016
24. Bursztein E, Benko B, Margolis D, Pietraszek T, Archer A, Aquino A, Pitsillidis A, Savage S (2014) Handcrafted fraud and extortion: manual account hijacking in the wild. In: Proceedings of the 2014 conference on internet measurement conference. ACM, New York, pp 347–358. doi:10.1145/2663716.2663749
25. PhishTank (2006) PhishTank | Join the fight against phishing. https://www.phishtank.com/. Accessed 4 Jan 2016
26. Han W, Cao Y, Bertino E, Yong J (2012) Using automated individual white-list to protect web digital identities. Expert Syst Appl 39(15):11861–11869. doi:10.1016/j.eswa.2012.02.020
27. Alsharnouby M, Alaca F, Chiasson S (2015) Why phishing still works: user strategies for combating phishing attacks. Int J Hum Comput Stud 82:69–82. doi:10.1016/j.ijhcs.2015.05.005
28. Sheng S, Holbrook M, Kumaraguru P, Cranor LF, Downs J (2010) Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, pp 373–382. doi:10.1145/1753326.1753383
29. Alseadoon I, Othman M, Chan T (2015) What is the influence of users' characteristics on their ability to detect phishing emails? Advanced computer and communication engineering technology. Springer International Publishing, Berlin, pp 949–962. doi:10.1007/978-3-319-07674-4_89
30. Trustworthy Internet Movement (2015) SSL pulse. https://www.trustworthyinternet.org/ssl-pulse/. Accessed 4 Jan 2016
31. Chen MC, Anderson JR, Sohn MH (2001) What can a mouse cursor tell us more? Correlation of eye/mouse movements on web browsing. In: Extended abstracts on human factors in computing systems (CHI). ACM, New York, pp 281–282. doi:10.1145/634067.634234
32. Felt A, Wagner D (2011) Phishing on mobile devices. In: Web 2.0 security and privacy workshop (W2SP)

33. Nurse JRC, Creese S, Goldsmith M, Lamberts K (2011) Trustworthy and effective communication of cybersecurity risks: a review. In: Proceedings of international workshop on socio-technical aspects in security and trust (STAST). IEEE, New York, pp 60–68. doi:10.1109/STAST.2011.6059257
34. Yang W, Chen J, Xiong A, Proctor RW, Li N (2015) Effectiveness of a phishing warning in field settings. In: Proceedings of the symposium and bootcamp on the science of security. ACM, New York, pp 14–1142. doi:10.1145/2746194.2746208
35. Zielinska OA, Tembe R, Hong KW, Ge X, Murphy-Hill E, Mayhorn CB (2014) One phish, two phish, how to avoid the internet phish analysis of training strategies to detect phishing emails. In: Proceedings of the human factors and ergonomics society annual meeting, vol 58. SAGE Publications, Thousand Oaks, pp 1466–1470. doi:10.1177/1541931214581306
36. Arachchilage NAG, Love S, Beznosov K (2016) Phishing threat avoidance behaviour: an empirical investigation. Comput Hum Behav 60:185–197. doi:10.1016/j.chb.2016.02.065
37. Nurse JRC, Creese S, Goldsmith M, Lamberts K (2011) Guidelines for usable cybersecurity: past and present. In: Proceedings of the third international workshop on cyberspace safety and security (CSS), IEEE, New York, pp 21–26. doi:10.1109/CSS.2011.6058566
38. Bravo-Lillo C, Cranor LF, Downs J, Komanduri S (2010) Bridging the gap in computer security warnings: a mental model approach. IEEE Secur Priv 2:18–26. doi:10.1109/MSP.2010.198