Why Phishing Works

Rachna Dhamija
rachna@deas.harvard.edu
Harvard University

J. D. Tygar
tygar@berkeley.edu
UC Berkeley

Marti Hearst
hearst@sims.berkeley.edu
UC Berkeley

ABSTRACT

To build systems shielding users from fraudulent (or *phishing*) websites, designers need to know which attack strategies work and why. This paper provides the first empirical evidence about which malicious strategies are successful at deceiving general users. We first analyzed a large set of captured phishing attacks and developed a set of hypotheses about why these strategies might work. We then assessed these hypotheses with a usability study in which 22 participants were shown 20 web sites and asked to determine which ones were fraudulent. We found that 23% of the participants did not look at browser-based cues such as the address bar, status bar and the security indicators, leading to incorrect choices 40% of the time. We also found that some visual deception attacks can fool even the most sophisticated users. These results illustrate that standard security indicators are not effective for a substantial fraction of users, and suggest that alternative approaches are needed.

Author Keywords

Security Usability, Phishing.

ACM Classification Keywords

H.1.2 [User/Machine Systems]: Software psychology; K.4.4 [Electronic Commerce]: Security.

Acknowledgements: Dr. Dhamija is currently at the Center for Research in Computation and Society at Harvard University. The authors thank the National Science Foundation (grants EIA-01225989, IIS-0205647, CNS-0325247), the US Postal Service, the UC Berkeley XLab, and the Harvard Center for Research in Computation and Society for partial financial support of this study. The opinions in this paper are those of the authors alone and do not necessarily reflect those of the funding sponsor or any government agency.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
CHI 2006, April 22–27, 2006, Montréal, Québec, Canada.
Copyright 2006 ACM 1-59593-178-3/06/0004...$5.00.

INTRODUCTION

What makes a web site credible? This question has been addressed extensively by researchers in computer-human interaction. This paper examines a twist on this question: what makes a *bogus* website credible? In the last two years, Internet users have seen the rapid expansion of a scourge on the Internet: *phishing*, the practice of directing users to fraudulent web sites. This question raises fascinating questions for user interface designers, because both phishers and anti-phishers do battle in user interface space. Successful phishers must not only present a high-credibility web presence to their victims; they must create a presence that is so impressive that it causes the victim to fail to recognize security measures installed in web browsers.

Data suggest that some phishing attacks have convinced up to 5% of their recipients to provide sensitive information to spoofed websites [21]. About two million users gave information to spoofed websites resulting in direct losses of $1.2 billion for U.S. banks and card issuers in 2003 [20].[1]

If we hope to design web browsers, websites, and other tools to shield users from such attacks, we need to understand which attack strategies are successful, and what proportion of users they fool. However, the literature is sparse on this topic.

This paper addresses the question of why phishing works. We analyzed a set of phishing attacks and developed a set of hypotheses about how users are deceived. We tested these hypotheses in a usability study: we showed 22 participants 20 web sites and asked them to determine which ones were fraudulent, and why. Our key findings are:

- Good phishing websites fooled 90% of participants.
- Existing anti-phishing browsing cues are ineffective. 23% of participants in our study did not look at the address bar, status bar, or the security indicators.
- On average, our participant group made mistakes on our test set 40% of the time.

[1]Over 16,000 unique phishing attack websites were reported to the Anti-Phishing Working Group in November 2005 [2].

Carbyne Ex. 2043, p.1 of 10
Apple v. Carbyne
IPR2024-00332

- Popup warnings about fraudulent certificates were ineffective: 15 out of 22 participants proceeded without hesitation when presented with warnings.
- Participants proved vulnerable across the board to phishing attacks. In our study, neither education, age, sex, previous experience, nor hours of computer use showed a statistically significant correlation with vulnerability to phishing.

RELATED WORK

Research on Online Trust

Researchers have developed models and guidelines on fostering online consumer trust [1, 4, 5, 8, 9, 10, 11, 15, 16, 18, 19, 23, 28]. Existing literature deals with trustworthiness of website content, website interface design and policies, and mechanisms to support customer relations. None of these papers consider that these indicators of trust may be spoofed and that the very same guidelines that are developed for legitimate organizations can also be adopted by phishers.

Empirical research in online trust includes a study of how manipulating seller feedback ratings can influence consumer trust in eBay merchants [4]. Fogg et al. conducted a number of large empirical studies on how users evaluate websites [10, 11] and developed guidelines for fostering credibility on websites, e.g., "Make it easy to verify the accuracy of the information on your site" [9].

User Studies of Browser Security and Phishing

Friedman et al. interviewed 72 individuals about web security and found that participants could not reliably determine whether a connection is secure. Participants were first asked to define and make non-verbal drawings of a secure connection. They were next shown four screen shots of a browser connecting to a website and were asked to state if the connection was secure or not secure and the rationale for their evaluation [14]. In a related study, Friedman et al. surveyed 72 people about their concerns about potential risks and harms of web use [13].

We are aware of two empirical user studies that specifically focus on phishing. Wu et al. conducted a user study to examine the impact of anti-phishing toolbars in preventing phishing attacks [29]. Their results show that even when toolbars were used to notify users of security concerns, users were tricked into providing information 34% of the time.

Jagatic et al. investigated how to improve the success of phishing attacks by using the social network of the victim to increase the credibility of phishing email [17]. In the study, the experimenters gathered data from the Internet to create a social network map of university students, and then used the map to create forged phishing email appearing to be from friends. 72% of users responded to the phishing email that was from a friend's spoofed address, while only 16% of users responded in the control group to phishing email from an unknown address.

ANALYSIS OF A PHISHING DATABASE

The Anti Phishing Working Group maintains a "Phishing Archive" describing phishing attacks dating back to September 2003 [3]. We performed a cognitive walkthrough on the approximately 200 sample attacks within this archive. (A cognitive walkthrough evaluates the steps required to perform a task and attempts to uncover mismatches between how users think about a task and how the user interface designer thinks about the task [27].) Our goal was to gather information about which strategies are used by attackers and to formulate hypotheses about how lay users would respond to these strategies.

Below we list the strategies, organized along three dimensions: lack of knowledge, visual deception, and lack of attention. To aid readers who are unfamiliar with the topic, Table 1 defines several security terms.

Certificate (digital certificate, public key certificate): uses a digital signature to bind together a public key with an identity. If the browser encounters a certificate that has not been signed by a trusted *certificate authority*, it issues a warning to the user. Some organizations create and sign their own *self-signed certificates*. If a browser encounters a self-signed certificate, it issues a warning and allows the user to decide whether to accept the certificate .

Certificate Authority (CA): an entity that issues *certificates* and attests that a public key belongs to a particular identity. A list of trusted CAs is stored in the browser. A certificate may be issued to a fraudulent website by a CA without a rigorous verification process.

HTTPS: Web browsers use "HTTPS", rather than "HTTP" as a prefix to the URL to indicate that HTTP is sent over *SSL/TLS*.

Secure Sockets Layer (SSL) and Transport Layer Security (TLS): cryptographic protocols used to provide authentication and secure communications over the Internet. SSL/TLS authenticates a server by verifying that the server holds a *certificate* that has been digitally signed by a trusted *certificate authority*. SSL/TLS also allows the client and server to agree on an encryption algorithm for securing communications.

Table 1: Security Terms and Definitions

1. Lack of Knowledge

1a) Lack of computer system knowledge. Many users lack the underlying knowledge of how operating systems, applications, email and the web work and how to distinguish among these. Phishing sites exploit this lack of knowledge in several ways. For example, some users do not understand the meaning or the syntax of domain names and cannot distinguish legitimate versus fraudulent URLs (e.g., they may think www.ebay-members-security.com belongs to www.ebay.com). Another attack strategy forges the email header; many users do not have the skills to distinguish forged from legitimate headers.

Carbyne Ex. 2043, p.2 of 10
Apple v. Carbyne
IPR2024-00332

1b) Lack of knowledge of security and security indicators. Many users do not understand security indicators. For example, many users do not know that a closed padlock icon in the browser indicates that the page they are viewing was delivered securely by SSL. Even if they understand the meaning of that icon, users can be fooled by its placement within the body of a web page (this confusion is not aided by the fact that competing browsers use different icons and place them in different parts of their display). More generally, users may not be aware that padlock icons appear in the browser "chrome" (the interface constructed by the browser around a web page, e.g., toolbars, windows, address bar, status bar) only under specific conditions (i.e., when SSL is used), while icons in the content of the web page can be placed there arbitrarily by designers (or by phishers) to induce trust.[2]

Attackers can also exploit users' lack of understanding of the verification process for SSL certificates. Most users do not know how to check SSL certificates in the browser or understand the information presented in a certificate. In one spoofing strategy, a rogue site displays a certificate authority's (CA) trust seal that links to a CA webpage. This webpage provides an English language description and verification of the legitimate site's certificate. Only the most informed and diligent users would know to check that the URL of the originating site and the legitimate site described by the CA match.

2. Visual Deception

Phishers use visual deception tricks to mimic legitimate text, images and windows. Even users with the knowledge described in (1) above may be deceived by these.

2a) Visually deceptive text. Users may be fooled by the syntax of a domain name in "typejacking" attacks, which substitute letters that may go unnoticed (e.g. www.paypai.com uses a lowercase "i" which looks similar to the letter "l", and www.paypa1.com substitutes the number "1" for the letter "l"). Phishers have also taken advantage of non-printing characters [25] and non-ASCII Unicode characters [26] in domain names.

2b) Images masking underlying text. One common technique used by phishers is to use an image of a legitimate hyperlink. The image itself serves as a hyperlink to a different, rogue site.

2c) Images mimicking windows. Phishers use images in the content of a web page that mimic browser windows or dialog windows. Because the image looks exactly like a real window, a user can be fooled unless he tries to move or resize the image.

2d) Windows masking underlying windows. A common phishing technique is to place an illegitimate browser window on top of, or next to, a legitimate window. If they have the same look and feel, users may mistakenly believe that both windows are from the same source, regardless of variations in address or security indicators. In the worst case, a user may not even notice that a second window exists (browsers that allow borderless pop-up windows aggravate the problem).

2e) Deceptive look and feel. If images and logos are copied perfectly, sometimes the only cues that are available to the user are the tone of the language, misspellings or other signs of unprofessional design. If the phishing site closely mimics the target site, the only cue to the user might be the type and quantity of requested personal information.

3. Bounded Attention

Even if users have the knowledge described in (1) above, and can detect visual deception described in (2) above they may still be deceived if they fail to notice security indicators (or their absence).

3a) Lack of attention to security indicators. Security is often a secondary goal. When users are focused on their primary tasks, they may not notice security indicators or read warning messages. The image-hyperlink spoof described in (2b) above would thwarted if user noticed the URL in the status bar did not match the hyperlink image, but this requires a high degree of attention. Users who know to look for an SSL closed-padlock icon may simply scan for the presence of a padlock icon regardless of position and thus be fooled by an icon appearing in the body of a web page.

3b) Lack of attention to the absence of security indicators. Users do not reliably notice the absence of a security indicator. The Firefox browser shows SSL protected pages with four indicators (see Figure 1). It shows none of these indicators for pages not protected by SSL. Many users do not notice the absence of an indicator, and it is sometimes possible to insert a spoofed image of an indicator where one does not exist.

STUDY: DISTINGUISHING LEGITIMATE WEBSITES

To assess the accuracy of the hypotheses resulting from our cognitive walkthrough of phishing sites, we conducted a usability study. We presented participants with websites that appear to belong to financial institutions and e-commerce companies, some spoofed and some real. The participants' task was to identify legitimate and fraudulent sites and describe the reasoning for their decisions.

[2]For user convenience, some legitimate organizations allow users to login from non-SSL pages. Although the user data may be transmitted securely, there is no visual cue in the browser to indicate if SSL is used for form submissions. To "remedy" this, designers resort to placing a padlock icon in the page content, a tactic that phishers also exploit.

Carbyne Ex. 2043, p.3 of 10
Apple v. Carbyne
IPR2024-00332

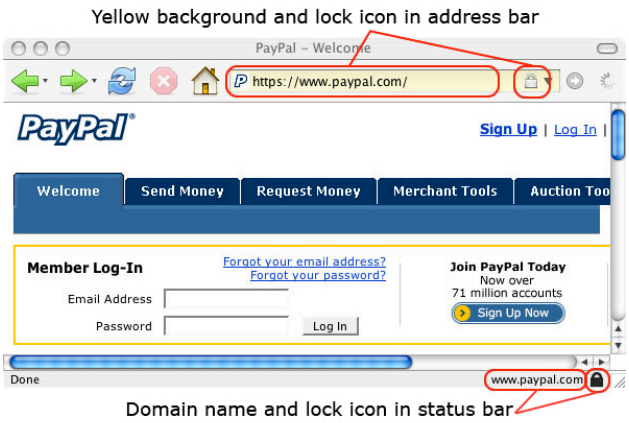

Figure 1: Visual Security Indicators in Mozilla Firefox Browser v1.0.1 for Mac OS X.

Our study primed participants to look for spoofs. Thus, these participants are likely better than "real-world" (unprimed) users at detecting fraudulent web sites. If our participants are fooled, real-world users are likely to also be fooled.

We focus on factors that are important for evaluating website security and authenticity, rather than the phishing email that lures users to those websites. (Several studies evaluate users' ability to detect fraudulent phishing email [17, 22]. As discussed in the related work section, there is less empirical data on how users verify the security and authenticity of potentially fraudulent websites.)

Collection and Selection of Phishing Websites

Using a web archiving application, we collected approximately 200 unique phishing websites, including all related links, images and web pages up to three levels deep for each site. To find these sites, we used phishing email that we and our colleagues received in June and July 2005. MailFrontier, an anti-spam firm, provided us additional phishing URLs harvested from phishing email received between July 20 and July 26, 2005.

We selected nine phishing attacks, representative in the types of targeted brands, the types of spoofing techniques, and the types of requested information. We also created three phishing websites, using advanced attacks observed by organizations monitoring phishing attacks [3, 24], but otherwise not represented in our sample. (Full descriptions of these sites are in [6].)

3.4.2 Study Design

We used a within-subjects design, where every participant saw every website, but in randomized order. Participants were seated in front of a computer in a University classroom and laboratory. We used an Apple G4 Powerbook laptop running MAC OS X (version 10.3.9). We used the Mozilla Firefox browser version 1.0.1 for Mac OS X. Firefox offers advanced security features (see Figure 1).

We created a webpage describing the study scenario and giving instructions, followed by a randomized list of hyperlinks to websites labeled "Website 1", "Website 2", etc. We intentionally did not label the hyperlinks with the name of the website or organization that was supposedly being linked to. Therefore, participants had no expectations about the site that they were about to visit or about upcoming sites they would visit next.

We presented participants with 20 websites; the first 19 were in random order:

- 7 legitimate websites
- 9 representative phishing websites
- 3 phishing websites constructed by us using additional phishing techniques
- 1 website requiring users to accept a self-signed SSL certificate (this website was presented last to segue into an interview about SSL and certificates).

Each website that we presented was fully functioning, with images, links and sub-pages that users could interact with as they normally would with any website. The archived phishing web pages were hosted on an Apache web server running on the computer that was used for the user study. The settings of the computer (i.e., hosts file, DNS settings, Apache configuration files) were modified so that the website appeared in the browser exactly as it did in the actual phishing attack, with the same website structure and same URL. To display the legitimate websites, we provided a hyperlink to the actual website.

Scenario and Procedure

We presented the following scenario to participants:

"Imagine that you receive an email message that asks you to click on one of the following links. Imagine that you decide to click on the link to see if it is a legitimate website or a "spoof" (a fraudulent copy of that website)."

We told participants that they could interact with the website as users usually would, that the websites were randomly selected, and that they might see multiple copies of the same website. We informed participants any website may be legitimate or not, independent of what they previously saw.

Participants signed a consent form, answered basic demographic questions, and read the study scenario and instructions. We then showed them the list of linked websites. As each website was viewed, we asked the participant to say if the site was legitimate or not, state their confidence in their evaluation (on a scale of 1-5) and their reasoning. Participants were encouraged to talk out loud and vocalize their decision process. We also asked par-

Carbyne Ex. 2043, p.4 of 10
Apple v. Carbyne
IPR2024-00332

ticipants if they had used this website in the past or if they had an account at the website's organization.

We also observed participants' interaction with a website that required accepting a self-signed SSL certificate. Afterwards, we asked participants about their knowledge and use of certificates and SSL. We also asked about experiences with phishing in general.

Finally, we provided a debriefing, where we educated the participants about the incorrect answers they had given. We provided a brief overview of domain names and SSL indicators and how to interpret them. We then revisited the other websites seen in the study to discuss the mistakes and correct assumptions that were made.

Participant Recruitment and Demographics

Our 22 participants were recruited by a university subjects recruiting service. This service uses a subscription based email list, which consists of students and staff who sign up voluntarily to participate in user studies. The only requirement was that participants be familiar with use of computers, email and the Web. They received a $15 fee for participating.

The participants were 45.5% male (10 participants) and 54.5% female (12 participants). Age ranged from 18 to 56 years old (M=29.9, s.d.= 10.8, variance=116).

Half of the participants were university staff, and half were students. 19 participants (86%) were in non-technical fields or areas of study. 3 (14%) were in technical fields. Of the staff, 8 participants (73%) had a Bachelors degree, 2 participants (18%) had a Masters degree, and 1 participant (9%) earned a J.D. degree. Of the students, 7 participants (63.6%) were Bachelors, 2 (18%) were Masters students, and 2 (18%) were Ph.D. students.

As their primary browser, 11 participants (50%) use Microsoft Internet Explorer, 7 (32%) use Mozilla Firefox , 2 (9%) reported using "Mozilla" (version unknown), and 1 (4.5%) uses Apple Safari. As their primary operating system, 13 participants (59%) use Windows XP, 6 (27%) use Mac OS X, 2 (9%) use Windows 2000, and 1 (4.5%) uses "Windows" (version unknown). Most participants regularly use more than one type of browser and operating system.

Hours of computer usage per week ranged from 10 to 135 hours (M=37.8, s.d.= 28.5, variance=810.9). 18 participants regularly use online banking (or another financial service such as online bill payment or Paypal). 20 participants said they regularly shop online.

RESULTS

Participant Scores and Behavior

The sum of the number of correctly identified legitimate and spoofed websites forms the participant's score. Scores ranged from 6 to 18 correctly identified websites, out of 19 websites. (Mean 11.6, s.d.=3.2, variance=10.1).

Sex: There is no significant difference when comparing the mean scores of males and females (t=1.97, df=20, p=.064). The mean score is 13 for males (s.d.=3.6, variance=13.1) and 10.5 for females (s.d.=2.3, variance=5.4).

Age: There is no significant correlation between participants' scores and ages (r=.065, df=20, p=.772). Younger participants did not perform better than older participants.

Education level: There is no significant correlation between education level and scores (*Spearman rho*=.24, n=22, p=.283).

Hours using the computer: There is no significant correlation between the weekly number of hours participants used computers and their scores (r= -.242, df=20, p=.278). One participant judged 18 out of 19 sites correctly but used computers only 14 hours per week while another participant judged only 7 websites correctly even though he spent 90 hours per week using computers.

Previous use of browser, operating system or website: There is no significant correlation between the score and the primary or secondary type of browser or operating systems used by participants. There is also no significant correlation between the previous use of a website and the ability to correctly judge the legitimacy of the purported (legitimate or phishing) website.

In summary, among our participants, we did not observe a relationship between scores and sex, age, educational level or experience. A larger study is needed to establish or rule out the existence of such effects in the general population.

Strategies for Determining Website Legitimacy

Participants used a variety of strategies to determine whether a site was legitimate or not. We categorized participants by the types and number of factors they used to make decisions, using their behavior, statements made while evaluating websites and answers to our questions. Participants' statements about indicators that they do or do not pay attention to were consistent with their behavior in the study.

Type 1: Security indicators in website content only

23% (5) participants used only the content of a webpage to determine legitimacy; including logos, layout and graphic design, presence of functioning links and images, types of information presented, language, and accuracy of information. As we discuss below, many participants always looked for a certain type of content (e.g.. a pad-

Carbyne Ex. 2043, p.5 of 10
Apple v. Carbyne
IPR2024-00332

lock icon, contact information, updated copyright information) in making their decision. None of these participants mentioned the address bar or any other part of the browser chrome as factors in their judgments. Later, each confirmed that they do not look at these regions of the browser. For example, one said, "I never look at the letters and numbers up there [in the address bar]. I'm not sure what they are supposed to say".

Participants in this category were at a disadvantage and received the five lowest scores (6, 7, 7, 9, 9). Without looking at the URL, they could not recognize the difference between two sites that looked similar but that were hosted on different servers. For example, when the phishing page linked to a privacy policy hosted on a legitimate site, this group of participants confused the legitimate and bogus sites. Phishers can exploit these users by creating what appears to be a rich and fully functioning website by linking to underlying pages on a legitimate site to boost credibility.

Type 2: Content and domain name only
36% (8) participants used the address bar to make judgments in addition to the content factors mentioned above. This set of participants did not look for or notice any SSL indicators (including "HTTPS" in the address bar). However, this category of users was at least aware when the address bar changed as they move from site to site. They were able to distinguish addresses that contain IP numbers from those that contained domain names. Many did not know what an IP address is (participants referred to it as a "redirector address", a "router number", "ISP number", "those number thingies in front of the name"), however, many of these participants had their suspicions heightened when they saw an IP address instead of a domain name.

Type 3: Content and address, plus HTTPS
9% (2) participants used the above factors but also relied on the presence of "HTTPS" in the address bar. These participants did not notice or look for the SSL padlock icon. In fact, one stated that she never noticed the SSL padlock icon in any browser chrome before this study (she was a Firefox and Safari user). The other participant did use "HTTPS" in the address bar as a judgment factor, but incorrectly stated that site icons (favicons) in address bars indicate authenticity better because they "cannot be copied".

Type 4: All of the above, plus padlock icon
23% (5) participants relied on all of the above factors, but also looked for or noticed a padlock icon in the browser chrome. In the interview, we discovered that even if they noticed the padlock icon, some participants gave more credence to padlock icons that appeared within the content of the page.

Type 5: All of above, plus certificates
9% (2) participants relied on all of the above factors and also checked the certificate that was presented to their browser in our study. Both said that they have checked certificates in the past and that they occasionally check them if they are uncertain about the site's identity (e.g., when the browser presents a warning).

Additional Strategies
Two participants in our study stated that in general, they would only question a website's legitimacy if more than the username and password was requested. One participant actually submitted her username and password to some websites in order to verify if it was a site at which she had an account. She stated that this is a strategy that she has used reliably in practice to determine site authenticity. Her reasoning was "What's the harm? Passwords are not dangerous to give out, like financial information is". This participant admitted she does use the same password for many sites, but never considered that passwords obtained at one website might be used for fraudulent purposes at another site. She used Type 1 strategies, with a score of 7 out of 19 of websites judged correctly.

Another participant was at the other end of the spectrum. He opened up another browser window, where he typed in all URLs by hand in order to compare these pages to every website presented in the study. He also occasionally used Yahoo to search for the organization in question. He would click on the top search result and compare it to the website presented in the study. He stated that ever since a family member was the victim of a PayPal phishing attack, he now follows these steps in practice to protect himself. He used Type 4 strategies and scored 18 out of 19 sites judged correctly.

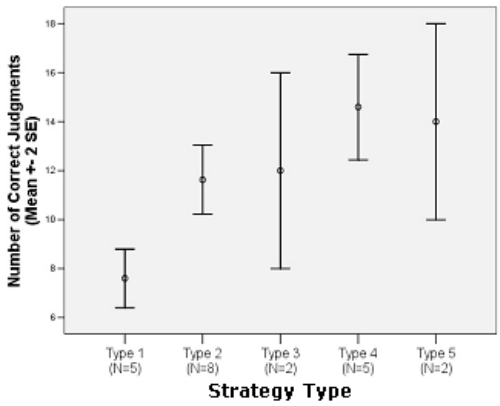

Figure 2: Mean Scores by Strategy Type (higher is better).

Comparison of Mean Scores Between Strategy Types
Figure 2 compares the mean number of websites judged correctly across strategy types. A one-way ANOVA re-

Carbyne Ex. 2043, p.6 of 10
Apple v. Carbyne
IPR2024-00332

veals that correct judgment scores differed significantly as a function of strategy ($F(4, 17) = 7.83$, $p = .001$). A Tukey post-hoc test reveals that the scores for Type 1 strategy users - those who use only the website content to determine legitimacy - differ significantly from Type 2, 4 and 5 strategy users.

Website Difficulty

After participants judged each website as legitimate or not, we asked them to rate how confident they were of this decision (on a scale of 1 to 5, where 1 was the least confident and 5 was the most confident). In general, participants were very confident of their decisions, whether they were correct or incorrect. The lowest average confidence level is 3.0.

Table 2 shows the websites ranked from most difficult (highest number of incorrect judgments) to least difficult (highest number of correct judgments), the average confidence level for each judgment, and the spoofing (fraudulent sites) or security (legitimate sites) strategies used.

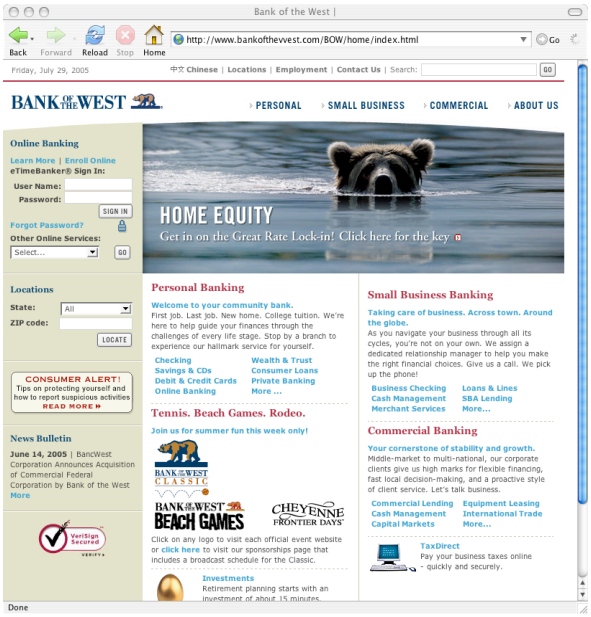

Figure 3: Bank of the West Phishing Site

Phishing Websites

Figure 3 shows the phishing website that fooled the most participants is an exact replica of the Bank of the West homepage. The website is hosted at "www.bankofthevvest.com", with two "v"s instead of a "w" in the domain name.

90.9% (20 participants) incorrectly judged this to be the legitimate Bank of the West website. 17 participants mentioned the content of the page as one reason for their decision. For many participants the "cute" design, the level of detail and the fact that the site does not ask for a great deal of information were the most convincing factors. Two participants mentioned the animated bear video that appears on the page, (e.g., "because that would take a lot of effort to copy"). Participants in general found this animation appealing and many reloaded the page just to see the animation again.

8 participants relied on links to other sites to support their decisions. 6 of these clicked on the Verisign logo: when clicked, a window popped up a window displaying an SSL protected webpage, hosted at Verisign, that shows the SSL certificate status of www.bankofthewest.com. Unfortunately, any site can provide a link to this popup page[3] in order to gain credibility. A participant must compare the URL displayed in the popup to the URL in the address bar to detect that they are not referring to the same website.

One participant clicked on a link to display the Chinese version of this website, which linked to the actual Bank of the West website. Her native language is Chinese, and she believed that a "fake website could never be this good". One participant clicked on a link to a "consumer alert", a link to the real website that describes tips for protecting against phishing attacks.

In fact, three participants said the correctness of the URL was the primary factor in deciding that this was a legitimate site. One of these was a Bank of the West account holder. He used a Type 5 strategy (i.e., used all browser security indicators) and had expert security knowledge. He stated that the use of "BOW/index.html" in the URL matched his memory of the legitimate web site. This indicates that even users who are knowledgeable and have familiarity with the website can be fooled by visual deception attacks.

9.1% (2 participants) correctly judged this to be a spoof site. Only one participant detected the double "v" in the domain name (she was the oldest participant in the 53-58 age range and used a Type 2 strategy with no security knowledge). One participant noticed an outdated date in the content of the webpage (many phishing sites display the date at which the page was copied from the original).

Participant Knowledge of Phishing and Security

We used participant responses to the websites to guide a semi-structured interview about their knowledge of browser security indicators and of phishing in general.

Knowledge and Experience with Phishing. 7 participants had never heard the term "phishing" before this study (some participants seemed genuinely surprised that these attacks even occur). However, all participants said that they do receive spam email that asks them to visit a web

[3] https://seal.verisign.com/splash?form_file=fdf/splash.fdf&dn=WWW.BANKOFTHEWEST.COM&lang=en

Carbyne Ex. 2043, p.7 of 10
Apple v. Carbyne
IPR2024-00332

Website	Real or Spoof	Phishing or Security Tactic Used (Partial List)	% Right (avg conf)	% Wrong (avg conf)
Bank Of the West	Spoof	URL (bankofthevvest.com), padlock in content, Verisign logo and certificate validation seal, consumer alert warning	9 (3.0)	91 (4.2)
PayPal	Spoof	Uses Mozilla XML User Interface Language (XUL) to simulate browser chrome w/ fake address bar, status bar and SSL indicators	18 (3.0)	81 (4.5)
Etrade	Real	3rd party URL (etrade.everypath.com), SSL, simple design, no graphics for mobile users	23 (4.6)	77 (4.2)
PayPal	Spoof	URL (paypal-signin03.com), padlock in content	41 (4.0)	59 (3.7)
PayPal	Spoof	URL (IP address), padlock in content	41 (3.9)	59 (4.5)
Capital One	Real	3rd party URL (cib.ibanking-services.com), SSL, dedicated login page, simple design	50 (3.9)	50 (3.5)
Paypal	Spoof	Screenshot of legitimate SSL protected Paypal page within a rogue webpage	50 (4.7)	50 (4.3)
Ameritrade	Spoof	URL (ameritrading.net)	50 (4.2)	50 (3.9)
Bank of America	Spoof	Rogue popup window on top of legitimate BOFA homepage, padlock in content	64 (4.2)	36 (4.4)
Bank Of The West	Spoof	URL (IP address), urgent anti-fraud warnings (requests large amount of personal data)	68 (4.8)	32 (4.4)
USBank	Spoof	URL (IP address), padlock in content, security warnings, identity verification (requests large amount of personal data)	68 (4.1)	32 (4.3)
Ebay	Spoof	URL (IP address), account verification (requests large amount of personal data)	68 (4.4)	32 (4.0)
Yahoo	Spoof	URL (center.yahoo-security.net), account verification (requests large amount of personal data)	77 (3.0)	23 (4.2)
NCUA	Spoof	URL (IP address), padlock in content, account verification (requests large amount of personal data)	82 (4.5)	18 (4.3)
Ebay	Real	SSL protected login page, TRUSTe logo	86 (4.4)	14 (4.0)
Bank Of America	Real	Login page on non-SSL homepage, padlock in content	86 (4.4)	14 (3.3)
Tele-Bears (Student Accounts)	Real	SSL protected login page	91 (4.7)	9 (4.5)
PayPal	Real	Login page on non-SSL homepage, padlock in content	91 (4.6)	9 (3.0)
Bank One	Real	Login page on non-SSL homepage, padlock in content	100 (4.0)	0 (N/A)

Table 2: Security or spoofing strategy employed by each site (spoof sites shown with white background, real sites gray).

site and provide information, log in to update accounts, verify records etc.

Participants mentioned various strategies for dealing with phishing email, including automatically deleting or filtering suspicious email (7 participants) and occasionally open suspicious email (15 participants). Of these, 6 do not ever click links from email and 5 will click on a link if it looks interesting. 2 only click on links if it from a website where they have an account. One will click on links from email only if associated with a specific transaction. One will click on any type of link at work where she has virus protection and system administrators to fix her machine, but never at home. All said they regularly click on links from friends, work colleagues and email sent by university organizations.

9 participants reported experiencing confusion about whether a site is legitimate. 5 reported serious incidents, where they or their family were tricked into proving personal information to a fraudulent online party.

Knowledge and Use of Padlock Icon and HTTPS. When asked about the meaning of the SSL padlock icon, 4 participants said they do not know what the padlock icon means and could not give any explanation for its presence. 5 participants mentioned the concept of security but could not identify what was secured or protected. 10 participants mentioned the concept of securing data that is sent by the user to the server. One stated that the SSL padlock icon indicated that the page sent from the server to the user was delivered securely. One participant incorrectly stated that it was a sign that the website could not read passwords or set cookies.

Carbyne Ex. 2043, p.8 of 10
Apple v. Carbyne
IPR2024-00332

15 participants stated that they never pay attention or notice the padlock icon, and that it does not affect their behavior (One of these participants, who was a Firefox and Safari participant, said that she had never noticed the security padlock in the browser before, only in the content of a page). 7 participants stated that they occasionally look for or notice the padlock.

13 participants stated that they never pay attention to "HTTPS" in the address bar. 5 stated that they never look at the address bar at all.

Knowledge and Use of Firefox SSL indicators. 17 participants did not notice the Firefox address bar SSL indicators during the study (3 of these were regular Firefox users and stated that they never noticed these indicators before they were pointed out in the study). Only 5 participants noticed the additional SSL indicators. Of these, 3 noticed the yellow background but didn't understand what it means (2 were Firefox users and one was an IE user). One Firefox user stated "I thought that was just part of the website design". Only 2 participants noticed the yellow background and understood that this was correlated with HTTPS (2 of these were Firefox users).

Knowledge and Use of Certificates. When presented with a browser warning for a self signed certificate, 15 participants immediately selected "OK" without reading the warning dialogue. The default option selected in this case is to "Accept this certificate temporarily for this session". When asked if they knew what they just accepted or declined, only one participant was able to correctly articulate the choice he had just made. 18 responded that they did not know and 3 gave incorrect answers (i.e., "I accepted the use of cookies"; "It asked me if I wanted to save my password on forms"; "It was a message from the website about spyware").

Only one participant gave a correct definition of the purpose of certificates and could interpret the website certificate that we asked the participants to inspect (he was a system administrator). 19 participants stated that they have never checked a certificate before. Only 3 participants stated that they ever checked a certificate (or saw something similar to the certificate that was shown).

Support for Hypotheses and Addition of New Ones

This study verified two types of hypotheses formulated in the examination of phishing sites. (The design of the study precludes testing for lack of attention, because we ask users to focus on security.) Participants made incorrect judgments because they lacked knowledge of how computer systems worked and did not have an understanding of security systems and indicators. More experienced participants were tripped up by visual deception, e.g., when the address was spoofed or when images of the browser chrome with security indicators were copied into website content. The study also revealed issues that we did not anticipate from the cognitive walkthrough:

1c) Lack of knowledge of web fraud. Some users don't know that spoofing websites is possible. Without awareness phishing is possible, some users simply do not question website legitimacy.

1d) Erroneous security knowledge. Some users have misconceptions about which website features indicate security. For example, participants assumed that if websites contained professional-looking images, animations, and ads, they assumed the sites were legitimate (influenced by well-known trust indicators, discussed below). Similarly, dedicated login pages from banks were less trusted than those originating from a homepage; several participants mentioned a lack of images and links as a reason for their distrust.

CONCLUSIONS

This study illustrates that even in the best case scenario, when users expect spoofs to be present and are motivated to discover them, many users cannot distinguish a legitimate website from a spoofed website. In our study, the best phishing site was able to fool more than 90% of participants.

Indicators that are designed to signal trustworthiness were not understood (or even noticed) by many participants. 5 out of 22 (23%) participants only used the content of the website to evaluate its authenticity, without looking at any other portions of the browser. A number of participants incorrectly said a padlock icon is more important when it is displayed within the page than if presented by the browser. Other participants were more persuaded by animated graphics, pictures, and design touches such as favicons (icons in the URL bar) than SSL indicators.

Furthermore, the indicators of trust presented by the browser are trivial to spoof. By using very simple spoofing attacks, such as copying images of browser chrome or the SSL indicators in the address bar or status bar, we were able to fool even our most careful and knowledgeable users.

Knowing this, phishers can falsify a rich and fully functioning site with images, links, logos and images of security indicators, and a significant fraction of our participants were confident that the spoofed websites were legitimate. Similarly, legitimate organizations that follow security precautions, such as allowing users to only login from dedicated SSL protected pages, are penalized and were judged by some of our participants to be less trustworthy. Legitimate organizations further confused our participants by hosting secure pages with third parties, where the domain name does not match the brand name.

Our study suggests that a different approach is needed in the design of security systems. Rather than approaching

Carbyne Ex. 2043, p.9 of 10
Apple v. Carbyne
IPR2024-00332

the problem solely from a traditional cryptography-based security framework (what can we secure?) a usable design must take into account what humans do well and what they do not do well. When building a system that is designed to resist spoofing, we must assume uniform graphic designs can be easily copied, and we must help the user to distinguish legitimate security indicators from those that have been spoofed. It is not sufficient for security indicators to appear only under trusted conditions- it is equally, if not more, important to alert users to the untrusted state. Finally, security interface designers must consider that indicators placed outside of the user's periphery or focus of attention (e.g., using colors in the address bar to indicate suspicious and trusted sites [12]) may be ignored entirely by some users.

To address some of these issues, we have designed and are currently testing a new approach that allows a remote server to prove its identity in a way that is easy for a user to verify (exploiting the human ability to easily match images) but difficult for an attacker to spoof [7].

REFERENCES

1. Ang, L., C. Dubelaar, & B. Lee. To Trust or Not to Trust? A Model of Internet Trust From the Customer's Point of View. *Proc. 14th Bled E-Commerce Conf.* (2001), 25-26.
2. Anti-Phishing Working Group. *Phishing Activity Trends Report November 2005* (2005).
3. Anti-Phishing Working Group Phishing Archive. http://anti-phishing.org/phishing_archive.htm
4. Ba, S. & P. Pavlov. Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior. *MIS Quarterly*, 26, 3 (2002), 243-268.
5. Cheskin Research. *E-commerce Trust Study* (1999).
6. Dhamija, R. Authentication for Humans: The Design and Analysis of Usable Security Systems. Ph.D. Thesis, University of California Berkeley (2005).
7. Dhamija, R. & J. D. Tygar. The Battle Against Phishing: Dynamic Security Skins. *Proc. SOUPS* (2005).
8. Egger, F.N. Affective Design of E-commerce User Interfaces: How to Maximize Perceived Trustworthiness. *Proc. Intl. Conf. Affective Human Factors Design* (2001), 317-324.
9. Fogg, B. J. Stanford Guidelines for Web Credibility. *Res. Sum. Stanford Persuasive Tech. Lab.* (2002).
10. Fogg, B. J. et al. How Do Users Evaluate the Credibility of Web Sites?: A Study with Over 2,500 Participants. *Proc. DUX* (2003).
11. Fogg, B. J. et al. What Makes Web Sites Credible?: A Report on a Large Quantitative Study. *Proc. CHI* (2001), 61-68.
12. Franco, R. *Better Website Identification and Extended Validation Certificates in IE7 and Other Browsers*. IEBlog, Nov. 21, 2005.
13. Friedman, B. et al. Users' Conceptions of Risks and Harms on the Web: A Comparative Study. *Ext. Abs. CHI* (2002), 614-615.
14. Friedman, B. et al. Users' Conceptions of Web Security: A Comparative Study. *Ext. Abs. CHI* (2002), 746-747.
15. Gefen, D. Reflections on the Dimensions of Trust and Trustworthiness Among Online Consumers. *ACM SIGMIS Database, 33, 3* (2002), 38-53.
16. Hemphill, T. Electronic Commerce and Consumer Privacy: Establishing Online Trust in the U.S. Digital Economy. *Bus. & Soc. Rev.*, 107, 2 (2002), 331-239.
17. Jagatic, T., N. Johnson, & M. Jakobsson. *Phishing Attacks Using Social Networks* (*Indiana U. Human Subject Study 05-9892 & 05-9893)*. (2005)
18. Kim, D., Y. Song, S. Braynov, & H. Rao. A B-to-C Trust Model for Online Exchange. *Proc. Americas Conf. on Inf. Sys*. (2001), 784-787.
19. Lee, M. & E. Turban. A Trust Model for Consumer Internet Shopping. *Intl J. Elec. Commerce*, 6, 1, (2001), 75-91.
20. Litan, A. *Phishing Attack Victims Likely Targets for Identity Theft*. Gartner Research (2004).
21. Loftesness, S. *Responding to "Phishing" Attacks*. Glenbrook Partners (2004).
22. MailFrontier, *MailFrontier Phishing IQ Test II* (2005).
23. Princeton Survey Research Associates, *A Matter of Trust*. (2002).
24. Secunia. http://secunia.com/.
25. Secunia, *Internet Explorer URL Spoofing Vulnerability* (2004).
26. Secunia, *Multiple Browsers Vulnerable to the IDN Spoofing Vulnerability* (2005).
27. Stone, D. et al. *User Interface Design & Evaluation*. Elsevier (2005).
28. Wang, Y & H. Emurian. An Overview of Online Trust. *Computers in Human Behavior*, 21, 1 (2005), 105-125.
29. Wu, M., R. Miller, & S. Garfinkel. Do Security Toolbars Actually Prevent Phishing Attacks? *Proc. CHI* (2006).

Carbyne Ex. 2043, p.10 of 10
Apple v. Carbyne
IPR2024-00332