

Do Security Toolbars Actually Prevent Phishing Attacks?

Min Wu, Robert C. Miller, Simson L. Garfinkel
 MIT Computer Science and Artificial Intelligence Lab
 32 Vassar Street, Cambridge, MA 02139
 {minwu, rcm, simsong}@csail.mit.edu

ABSTRACT

Security toolbars in a web browser show security-related information about a website to help users detect phishing attacks. Because the toolbars are designed for humans to use, they should be evaluated for usability – that is, whether these toolbars really prevent users from being tricked into providing personal information. We conducted two user studies of three security toolbars and other browser security indicators and found them all ineffective at preventing phishing attacks. Even though subjects were asked to pay attention to the toolbar, many failed to look at it; others disregarded or explained away the toolbars' warnings if the content of web pages looked legitimate. We found that many subjects do not understand phishing attacks or realize how sophisticated such attacks can be.

Author Keywords

World Wide Web and Hypermedia, E-Commerce, User Study, User Interface Design.

ACM Classification Keywords

H.5.2 User Interfaces, H.1.2 User/Machine Systems, D.4.6 Security and Protection.

INTRODUCTION

Phishing has become a significant threat to Internet users. Phishing attacks typically use legitimate-looking but fake emails and websites to deceive users into disclosing personal or financial information to the attacker. Users can also be tricked into downloading and installing hostile software, which searches the user's computer or monitors online activities to steal private information.

Phishing attacks are on the rise. According to the Anti-Phishing Working Group (APWG), 2870 phishing sites appeared in March 2005, a 28% increase *per month* since July 2004. [2] A survey sponsored by TRUSTe found 70% of the respondents had visited a phishing site; over 15%

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.

SpoofStick

You're on **paypal.com**

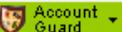
Netcraft Toolbar

Since: [Oct 2001](#) Rank: [41](#) [Site Report](#)  [US] [eBay, Inc](#)

TrustBar

 Identified by  The Value of Trust™

eBay Account Guard

 Search 

SpoofGuard

 www.paypal.com

Figure 1. Existing security toolbars

admitted to having provided personal data to a phishing site; and US consumers have lost an estimated \$500 million as a result of these attacks. [15]

APWG has collected and archived many phishing attacks. A typical example is an attack against eBay customers, first reported in March 2004. [1] The attack starts with an email claiming that the recipient's account information is invalid and needs to be updated by visiting the provided link. The message appears to come from `S-Harbor@eBay.com`, and the link apparently points to `cgil.ebay.com`, but actually leads to `210.93.131.250`, a server in South Korea with no relationship to eBay. Following the link produces a web page that looks legitimate, with an eBay logo and page design, and asks for the victim's credit card, Social Security number, eBay username and password. Clicking the submit button sends the data to the hostile server, where it is collected and used by the attackers.

Many proposals for stopping phishing attacks rely on a *security toolbar* that displays warnings or security-related information in the web browser's interface. Figure 1 shows some existing security toolbars:

- SpoofStick [20] displays the website's real domain name, in order to expose phishing sites that obscure their domain name. An attack might use a legitimate-looking domain name as a sub-domain, *e.g.*, `www.paypal.com.wws2.us` to fool users; SpoofStick would display this domain as `wws2.us`.

- Netcraft Toolbar [16] displays information about the site, including the domain's registration date, hosting country, and popularity among other toolbar users. This information is thought to be helpful in detecting phishing sites because most phishing sites are short-lived compared to the legitimate sites they imitate, and a large number of phishing sites spoof US-based corporations but are registered in other countries.
- Trustbar [13] makes secure web connections (SSL) more visible by displaying the logos of the website and its certificate authority (CA). This is useful against phishing because many legitimate websites use SSL to encrypt the user's sensitive data transmission, but most phishing sites do not. Attackers avoid SSL because obtaining an SSL certificate from a well-known CA, such as VeriSign, requires site identity information that can be traced, and because using a CA that is not known to the browser will trigger a warning and thus might raise the user's suspicion.
- eBay's Account Guard [7] shows a green icon to indicate that the current site belongs to eBay or PayPal, a red icon to indicate a known phishing site found on a blacklist maintained by eBay, and a gray icon for all other sites.
- SpoofGuard [5] calculates a *spoof score* for the current web page using a set of heuristics derived from previous phishing attacks. It then translates this score into a traffic light: red for spoof scores above a threshold, indicating the page is probably hostile; yellow for scores in the middle; and green for low scores, indicating the page is probably safe.

In addition to these toolbars, existing browser indicators can also help users to detect phishing attacks. For example, the address bar displays the URL of the current web page, and the status bar displays a lock icon to indicate if the page was downloaded with SSL. To further differentiate SSL-downloaded pages, Mozilla Firefox changes the address bar's background from white to yellow and adds a lock icon in the address bar. Users are commonly advised by online security tips to pay attention to these kinds of indicators whenever they access a web site. [9]

There are several potential drawbacks to the security-toolbar approach:

- A toolbar is a small display in the peripheral area of the browser, compared to the large main window that displays the web content. Users may not pay enough attention to the toolbar at the right times to notice an attack.
- A security toolbar shows security-related information, but security is rarely the user's primary goal in web browsing. Users may not care about the toolbar's display even if they do notice it.

- If a toolbar sometimes makes mistakes and identifies legitimate sites as phishing sites, users may learn to distrust the toolbar. Then, when the toolbar correctly identifies a phishing site, the user may not believe it.

This paper describes two user studies we performed to find out why users get fooled by phishing attacks, to determine which attacks were more effective than others, and to evaluate the security toolbar approach for fighting phishing.

The rest of this paper is organized as follows. We begin by surveying related work. Next we discuss how we designed a user study to evaluate the security toolbars, with particular attention to the issues and tradeoffs of a study that intentionally attacks users. We present the results from this study and discuss why the security toolbars do not work as expected. The next section presents a follow-up study using the same methodology that supports our reasoning about why phishing attacks are effective. We conclude with some design principles for anti-phishing solutions based on the results and observations from our two user studies.

RELATED WORK

A growing number of user studies are investigating why phishing attacks are so effective against computer users.

Anti-spam firm MailFrontier Inc did a web survey on how well people can distinguish phishing emails from legitimate ones. [21] Subjects saw screenshots of ten emails but could not interact with them. About 28% of the time, subjects incorrectly identified the phishing emails as legitimate.

In April 2004, a study in London found that 34% of the respondents would give the researchers their password in exchange for a bar of chocolate. [4] The researchers did not test the passwords to see if they were accurate, however.

In April 2005, a study at Indiana University Bloomington showed that social context can make phishing attacks far more effective. [14] The researchers sent out phishing emails to university students, claiming to be from a friend, having mined friendship relations from a social networking site used on campus. The email led to a phishing site that asked for the subject's university username and password. 72% of the subjects provided valid usernames and passwords to the phishing site.

Whalen and Inkpen used an eye-tracker to study the user's attention to browser security indicators when doing secure online transactions. [22] Their study found that subjects often looked at the lock icon in the status bar, but rarely clicked on the lock and thus didn't learn anything about the site's certificate. By contrast, subjects in our studies rated the status bar least effective at preventing phishing attacks. We think that the difference is due to the fact that Whalen and Inkpen's subjects were explicitly told to pay attention to the security indicators in the browser, while our subjects were asked to detect fake websites, so the address bar and the URL were more useful indicators.

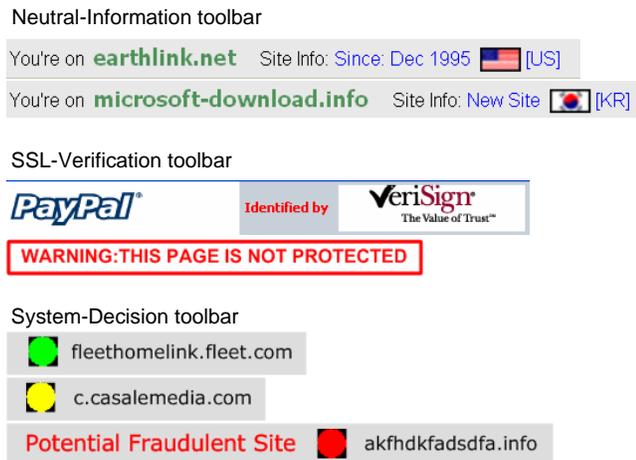


Figure 2. The three simulated toolbars tested in the study

At least two organizations have initiated phishing attacks against their own members, with the goal of teaching them to protect themselves. [3] The US Military Academy at West Point found that more than 80% of its cadets succumbed to a phishing attack by a fictional colonel. The State of New York mounted two attacks on its 10,000 employees; 15% were spoofed by the first attack, but only 8% by the second, which came three months later.

Besides the security toolbars we tested, there are other anti-phishing solutions that help users to differentiate the legitimate web sites from the phishing ones. Dynamic Security Skins [6] proposes to use a randomly generated visual hash to customize the browser window or web form elements to indicate the successfully authenticated sites. PassMark [18] includes a personalized image in a web page to indicate that the user has set up an account with the site. Google Safe Browsing for Firefox [12] pops up an alert when a user is on a web page that Google determines to be illegitimate. The content of the phishing page is also darkened to make it less convincing. Internet Explorer 7 [19] protects against phishing with a dynamically-updated black list of known phishing web sites, a client-side list of acceptable sites, and a set of heuristics. It blocks the user's activity with a detected phishing site. IE7 also has stricter enforcement of SSL certificates, in that it will not display websites with certificates that are invalid. A comprehensive survey of anti-phishing solutions can be found in [8].

STUDY DESIGN

To simplify the study design, we grouped the features of the five existing toolbars into three simulated toolbars (figure 2), based on the three types of information that existing security toolbars display:

The *Neutral Information toolbar* shows website information, such as domain name, hostname, registration date and hosting country, as SpoofStick and Netcraft Toolbar do. With this information, users must use their own judgment and experience to decide whether a site is legitimate or phishing.

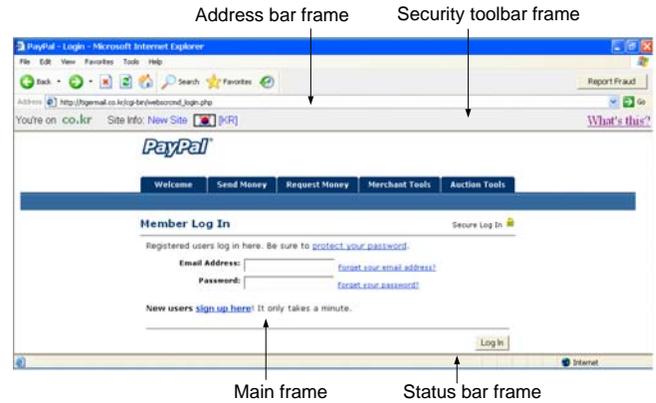


Figure 3. Browser simulation using HTML frames

The *SSL-Verification toolbar* differentiates sites that use SSL from those that do not. SSL sites are displayed with the site's logo and CA; a general warning message is displayed for other sites. This approach that imitates Trustbar seeks to make the user suspicious when a non-SSL page asks for sensitive information such as a password or credit card number.

The *System-Decision toolbar* displays a red light and the message "Potential Fraudulent Site" if it decides that a web page is actually a phishing attack, an approach that is similar in design to both eBay Account Guard and SpoofGuard. This display is easy for a user to interpret, but it requires the user to trust the toolbar's decision process, which is generally hidden from the user.

Study Implementation

In order to simulate attacks against users, we needed to completely control the display of the toolbars and other security indicators. Users in the study interacted with a simulated Internet Explorer built inside an HTML application running in full screen mode (figure 3). Different HTML frames displayed different browser components, including the security toolbars. The locations and sizes of the toolbars were consistent with the existing toolbars that they are based on. The Neutral-Information toolbar and the System-Decision toolbar were located below the address bar and above the main browsing window. The SSL-Verification toolbar was located below the title bar and above the menu bar. The address bar took the FireFox approach by using the yellow background and a lock icon to indicate SSL connections. The status bar also displayed a lock icon for SSL connections.

Our study simulated ideal phishing attacks whose content is a perfect copy of the actual website. This is realistic, since an attacker might not bother mirroring the entire site, but might simply act as a man-in-the-middle between the user and the real site. The attackers would pass the real web pages to the user and the user's submitted data to the real site and in the meantime capture the user's sensitive data during the online transaction. As such, the main frame in our browser always connected to the *real* website,

regardless of whether the site was supposed to be phishing or not. To simulate phishing attacks, we changed the appearance of the HTML frames that displayed the browser's security indicators—including the security toolbar, the address bar and the status bar—to indicate that the web page was served by an unusual source, e.g., `tigermail.co.kr` rather than `paypal.com`.

Study Scenario

Phishing is an attack that directly targets the human being in the security system. Simulating these kinds of attacks for study purposes raises some special problems. Chief among them is the *secondary goal* property articulated by Whitten and Tygar: in real life, security is rarely a user's primary goal. [23] The user is primarily concerned with other tasks, such as reading mail, buying a book, or editing a document. Avoiding disclosure of passwords or personal information may be important, but it isn't foremost in the user's mind.

In order to produce generalizable results, a lab study must be designed to preserve this behavior as much as possible. If we simply asked subjects to "identify the fake web pages," security would become their primary goal and hence lead them to pay attention and take precautions that they would be unlikely to take in real life.

We addressed this problem by creating a scenario which gave the subjects tasks to attend to other than security. With the given tasks, the subjects have to protect a secret from the attack. For ethical reasons we could not use actual financial data belonging to the subject. On the other hand, fake financial data cannot go through a real web site. One approach would be to use specially-crafted test bank account or credit card numbers provided by a financial institution. Another approach is to mirror the results of e-commerce transactions, as Whalen and Inkpen did. [22] A third approach is to set up our own fictional bank site, with which subjects could interact using the fake data. This approach is problematical for lab studies, since subjects would need some time to get used to the testing site and its transaction procedures.

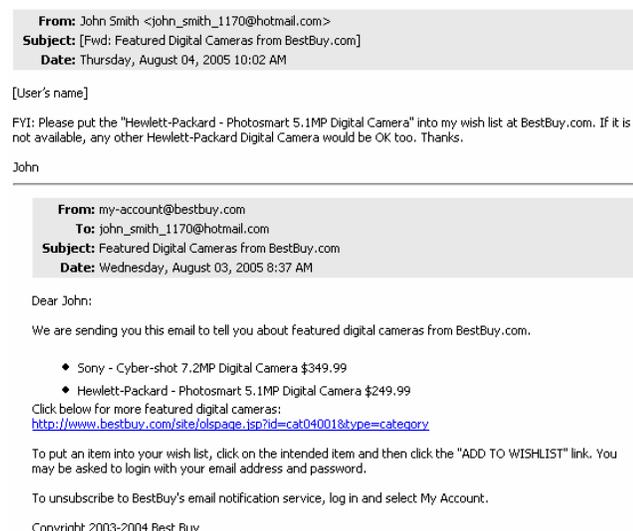


Figure 4. A sample email in the user study

Our studies took a different approach. We set up dummy accounts in the name of "John Smith" at various legitimate e-commerce websites and then asked the subjects to protect those passwords. With this approach, we could study a wide variety of existing websites with little setup. The subject played the role of John Smith's personal assistant and was given a printout of John's profile, including his fictitious personal and financial information and a list of his usernames and passwords. The task was to process 20 email messages, most of which were requests by John to handle a forwarded message from an e-commerce site. Each message contained a link for the user to click. Figure 4 shows a sample message.

Simulating Phishing Attacks

Five of the 20 forwarded emails were attacks, with links directing the users to a simulated phishing website. Each of these attacks represents a real phishing attack technique that has been recorded by APWG:

- *Similar-name attack:* Since one way that users authenticate web sites is by examining the URL displayed in the address bar, attackers can use a hostname that bears a superficial similarity to the imitated site's hostname. For example, we used `www.bestbuy.com.wv2.us` to spoof `bestbuy.com`.
- *IP-address attack:* Another way to obscure a server's identity is to display it as an IP address, e.g., `http://212.85.153.6/` to spoof `bestbuy.com`.
- *Hijacked-server attack:* Attackers sometimes hijack a server at a legitimate company and then use the server to host phishing attacks. For example, we used a hijacked site `www.btinternet.com` to spoof `bestbuy.com`.
- *Popup-window attack:* A popup-window attack displays the real site in the browser but puts a borderless window from the phishing site on top to request the user's personal information. Our phishing site displayed the `hollywoodvideo.com` site in the browser but popped up a window requesting the username and password. Although this pop-up window lacked an address bar and status bar, it nevertheless included the security toolbar.
- *PayPal attack:* The email message warns that John's account has been misused and needs to be reactivated, and points to a phishing website with hostname `tigermail.co.kr`. Unlike the other attacks, which simulate man-in-the-middle behavior while displaying the real web site, this attack requests not only a PayPal username and password, but also credit card and bank account information.

We consider the PayPal attack different from the other four attacks, which we call *wish-list attacks* because they merely asked the user to log in and modify a wish-list. First, the PayPal attack is like current phishing attacks that target online banks and financial services; the wish-list attacks target online retailers instead, which is not as common today, although growing. [10] The PayPal attack is greedy,

asking for lots of sensitive information; the wish-list attacks can only steal usernames and passwords. The PayPal attack is far more intimidating, urging users to reactivate their account and threatening to suspend their account if they did not do so immediately. We expected experienced Internet users to be more suspicious of the Paypal attack.

All three toolbars were configured to differentiate the legitimate sites from the phishing sites. None of the phishing sites used SSL so that the SSL-Verification toolbar always displayed a warning on them. On the System-Decision toolbar, all legitimate sites were displayed as trustworthy (green) but all the phishing sites were displayed as phishing (red) or unsure (yellow). On the Neutral-Information toolbar, the phishing sites and hijacked servers displayed as a “New Site” and some of them were displayed as they were hosted in other countries outside the US.

Toolbar Tutorial

Another question in this study design is when and how to give users a tutorial about the security toolbar. Few users read documentation in the real world; some may read an introduction when they download and install a security toolbar, but others may not read anything at all, particularly if the security features are bundled with the web browser.

Our pilot study found that the presence or absence of a tutorial has a strong effect on performance. When five pilot subjects received a printed tutorial explaining the security toolbar, showing how it looked for both legitimate and phishing websites, only 1 out of 15 attacks (7%) was successful. Another six pilot subjects received no printed tutorial; instead, we added a “What’s this?” link in each toolbar which displayed the tutorial in a popup window. These subjects succumbed to 17 out of 18 attacks (94%); not one subject one clicked the “What’s this?” link.

This result was problematic. In the former case, the printed tutorial gave the pilot subjects too strong a clue that security was the primary goal in the study. In the latter case, subjects had no idea what the security toolbar meant, or its role in preventing phishing attacks.

Based on this experience, we introduced the tutorial as part of the scenario. In the experiment, John Smith forwards to the subject an email from his company’s system administrator. The email says that a security toolbar has been installed on the company’s computers to prevent phishing attacks. The message contains a link to the tutorial. When John Smith forwarded this email to the subject, he explicitly requests that they be careful with his personal information.

The tutorial email appeared in the middle of the study, as the 11th of the 20 emails, where it could serve as a control to see how users behaved before and after seeing the tutorial. The PayPal attack was the 10th email because of its uniqueness. The remaining four attacks occurred at the 5th, 8th, 16th and 19th emails, with each type of wish-list attack

randomly assigned to one of these four positions. These fixed positions were chosen to space out the attacks.

Study Hypotheses

We define the *spooof rate* as the fraction of simulated attacks that successfully obtain John’s username and password or other sensitive information without raising the subject’s suspicion. We made two hypotheses: (1) that the spooof rates of all three toolbars would be substantially greater than 0, so that none of the toolbars effectively prevents attacks; and (2) that some toolbars would have better spooof rates than others. In particular, we expected that the System-Decision toolbar would have a lower spooof rate than the others because it used a simple traffic light metaphor, and these lights were always correct (at least in our simulation).

RESULTS AND DISCUSSION

A total of 30 subjects with previous experience in online shopping, 14 females and 16 males, were recruited by online and poster advertising at a college campus. Twenty subjects were college students from 10 different majors. All subjects had at least a college education. The average age was 27 (the range, 18 to 50). Each of the three security toolbars was tested on 10 subjects.

To gauge subjects’ experience with online shopping, we asked them which of our 19 selected e-commerce sites they had visited. All 30 subjects had used Amazon, and 25 or more had used PayPal, Travelocity, Bestbuy, and Yahoo. On average, each subject had used 10 of the sites in our study.

Before the study, subjects were given a consent form which (1) explained that the purpose of the study was to test web browser security indicators that detect fake web pages that look like pages from well-known legitimate websites; (2) indicated that the purpose of these fake websites is to trick people into making dangerous decisions or taking dangerous actions; and (3) encouraged the subjects to detect all the fake web pages and report them by clicking the “report fraud” button in the browser’s toolbar. All the subjects were required to read the consent form carefully, especially the study procedure part.

After the consent form, the subject was briefed about the John Smith scenario and their role as John Smith’s assistant. This briefing did not mention security at all.

We personally observed the subjects’ browsing behaviors during the study. We did not interrupt the study except when subjects clicked the “report fraud” button, at which point we asked them to explain why they reported fraud.

Security Awareness

One of the risks of using an artificial scenario is that users may not care about the fictional John Smith’s security at all. Fortunately, a number of indicators showed our subjects were behaving as if they did care about the security of John Smith’s accounts. Designing these kinds of secondary indicators into a security study turned out to be a good idea.

For example, 18 subjects unchecked the “Remember Me” checkbox on the login page of at least one site. This checkbox, which is generally checked by default and must be explicitly unchecked, controls whether John Smith’s login information is recorded in a cookie. Furthermore, 13 subjects explicitly logged out or tried to log out of at least one web site after finishing a task. These cautious subjects (23 in all) were protective of John Smith’s online identity and did not want the browser to remember the login sessions. We never told them anything about unchecking “Remember Me” or logging out. The other 7 subjects also exhibited suspicion and caution at least once in the study, either by reporting fraud or by trying to carefully explore a web site to determine if it was legitimate.

Subjects also demonstrated caution with *false alarms*—believing that a good site was an attack. Subjects did not finish tasks at good sites 3.3% of the time (13 out of 390 tasks) because of security concerns. There were six false alarms before the tutorial and seven after the tutorial. False alarms were generally due to browser warnings generated by the legitimate site (such as “You are about to be redirected to a connection that is not secure”).

The Wish-list Attacks

Figure 5 shows the spoof rates of wish-list attacks for each toolbar. These spoof rates, 45% for the Neutral-Information toolbar, 38% for the SSL-Verification toolbar, and 33% for the System-Decision toolbar, are all significantly higher than 0%, the ideal. No significant difference was found between the toolbars by a one-way ANOVA test. But this hardly matters since all the toolbars have high spoof rates.

Among the 30 subjects, 20 were spoofed by at least one wish-list attack (7 used the Neutral-Information toolbar, 6 used the SSL-Verification toolbar, and 7 used the System-Decision toolbar). We interviewed these subjects to find out why they did not recognize the attacks:

- 17 subjects (85%) mentioned in the interview that the web content looked professional or similar to what they had seen before. They were correct because the content *was* the real web site, but a high-quality phishing attack or man-in-the-middle can look exactly like the targeted

website as well. Seven of these subjects were observed to use security-related links *on the site itself* to decide if a site was legitimate or not—for example, clicking on the Verisign seal, the site’s privacy policy, contact information, copyright information, or a credit card security claim. Of course, attackers can and do fake these indicators. A lot of research has done to improve web credibility (*e.g.*, [11]), and attackers have clearly adopted these techniques.

- 12 subjects (60%) used rationalizations to justify the indicators of the attacks that they experienced. Nine subjects explained away odd URLs with comments like:

www.ssl-yahoo.com is a subdirectory of Yahoo!, like mail.yahoo.com.

sign.travelocity.com.zaga-zaga.us must be an outsourcing site for travelocity.com.

Sometimes the company [Target] has to register a different name [www.mytargets.com] from its brand. What if target.com has already been taken by another company?

Sometimes I go to a website and the site directs me to another address which is different from the one that I have typed.

I have been to other sites that used IP addresses [instead of domain names].

Four subjects explained away the popup window that asked for a username and password. One subject commented that she must have triggered the popup window herself, by clicking “Register for new account” instead of “Sign in for existing account”.

One subject explained away a toolbar message showing that Yahoo! was a “New Site” and located in Brazil by reasoning that Yahoo must have a branch in Brazil. Another explained away the warning on the System-Decision toolbar by saying that it was triggered because the web content is “informal,” just like a spam filter says that “this email is probably a spam.”

- Nine subjects (45%) said that the reason they were spoofed was that they were focused on finishing the study tasks—*i.e.*, dealing with John Smith’s email requests. Three explicitly mentioned that, although they noticed the security warnings, they had to take some risks to get the job done. Simply warning these subjects that something is wrong was not sufficient: they needed to be provided with a safe alternative way to achieve their goals.
- Five subjects (25%) claimed that they did not notice the toolbar display at all for some attacks.
- One subject extensively clicked links on the web pages to test whether the web site worked properly. By relying on the site’s behavior as an indication of its authenticity, this subject was fooled by *all* of the wish-list attacks.

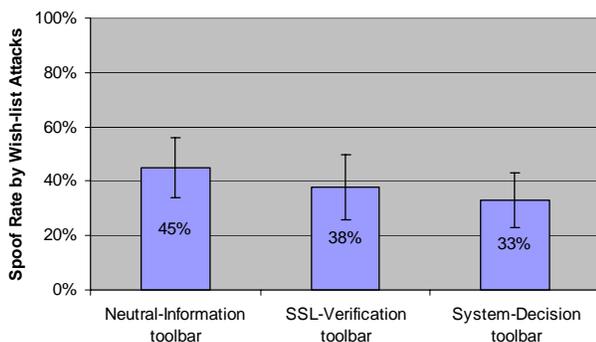


Figure 5. Spoof rates with different toolbars

The similar-name attack had the highest spoof rate, 50%, among all simulated phishing attack techniques. But no matter how the phishing URL is presented, the spoof rates are always high, with 43% for the hijacked-server attack and 33% for the IP-address attack. The popup-window attack had a relatively low spoof rate of 27%—many subjects thought that using the popup window for login information was abnormal and suspicious. The spoof rate differences were not significant by a one-way ANOVA test.

Learning Effects

Seven subjects clicked the toolbar’s “What’s this?” link before the tutorial email (1 using the Neutral-Information toolbar, 2 using the System-Decision toolbar, and 4 using the SSL-Verification toolbar.) Subjects using the SSL-Verification toolbar clicked “What’s this?” even before the first attack. We believe that this is because the toolbar displayed a warning message on all pages that did not use SSL, which is the case for many web pages.

We found that the difference in spoof rates for wish-list attacks before and after the subjects saw the tutorial, either by clicking the “What’s this?” link or by reading the tutorial email, to be statistically significant (one-tail $t(43) = 2.27$, $p = 0.014$). Figure 6 shows the spoof rate before the tutorial was 52%, while after the tutorial it dropped to 26%. Although a decrease was found with all three toolbars, the decrease was significant for the Neutral-Information toolbar (one-tail $t(18) = 1.84$, $p = 0.04$), marginally significant for the System-Decision toolbar (one-tail $t(12) = 1.52$, $p = 0.077$), and not significant for the SSL-Verification toolbar (one-tail $t(7) = 0.61$, $p = 0.28$).

Several subjects mentioned that the tutorial email helped them to pay more attention to the security toolbar and better understand its display, explaining the drop in spoof rate following the tutorial.

Subjects using the Neutral-Information toolbar and the System-Decision toolbar saw their spoof rates significantly drop following the tutorial. This was not true of subjects using the SSL-Verification toolbar. One explanation is that the toolbars had different levels of accuracy. We tried to make every toolbar accurate enough to distinguish phishing sites from legitimate sites. The System-Decision toolbar displayed a red or yellow light at the phishing sites but a green light at the good sites. The Neutral-Information

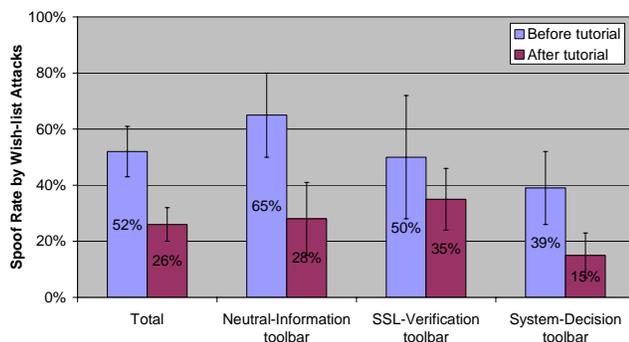


Figure 6. Spoof rates before and after the tutorial

toolbar showed all phishing sites as either a “new site” or hosted in a non-US country (or both), but all good sites as hosted in the US and in existence for several years. But it turned out that 9 of the 18 online stores that we chose for this study had login pages that were not protected by SSL, so the SSL-Verification toolbar produced warnings even for legitimate sites. Thus, the SSL-Verification toolbar failed to adequately distinguish fake sites from good ones.

The tutorial was not the only factor affecting subjects’ learning, of course. Another contribution to the decrease in the spoof rate before and after the tutorial is that the spoof rate steadily decreased for each attack as the subjects experienced more wish-list attacks and learned how to detect them, as shown in figure 7.

The PayPal Attack

As discussed above, the PayPal attack is very different from the wish-list attacks. The difference is reflected in the study. The PayPal attack had a significantly lower spoof rate (17%) than the wish-list attacks (38%) (two-tail $t(56) = -2.63$, $p = 0.01$). Ten subjects said that they had seen similar phishing emails in the real world, so they could detect the PayPal attack just by reading the email message, without even clicking through to the phishing site. The wish-list attacks have a lower spoof rate (28%) on these 10 subjects than the other 20 subjects (44%). But the difference is not significant (one-tail $t(23) = -1.36$, $p = 0.09$). Some subjects did not feel comfortable providing John Smith’s credit card and bank account information, and eventually noticed the suspicious signs from the toolbar or the suspicious URL from the address bar and thus avoided the attack.

However, there were still five subjects out of 30 (17%) who were tricked by the PayPal attack (at least one using each toolbar). Four were PayPal users in real life. They were spoofed because the content of the site looked authentic. One typical comment was “I’ve used PayPal before and this site looks exactly the same. If I trust a site from my experience, I am not suspicious.” They also justified the request as being reasonable. One subject said that “they need this information [the credit card and the bank account information] to charge me.” Thus, familiar phishing attacks can continue to be persuasive and effective, even with security toolbars to warn the user.

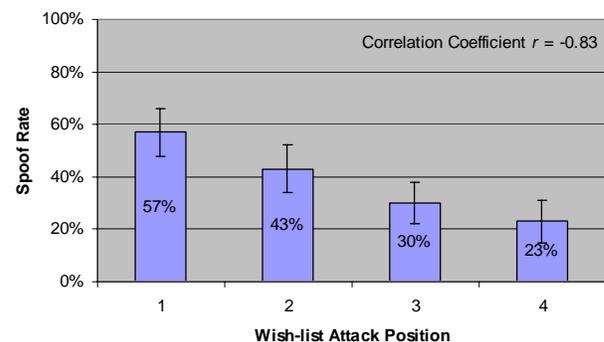


Figure 7. Spoof rates for wish-list attacks at different attack positions

Subjective Ratings and Comments on Toolbars

Subjects were asked at the conclusion of the study to rate the effectiveness of the address bar, status bar, the security toolbar that they used in differentiating authentic web sites from phishing sites, on a scale from -2 (very ineffective) to 2 (very effective). Figure 8 shows the mean ratings.

Among the three toolbars, the SSL-Verification toolbar was rated as less effective, although the difference was not significant. One reason might be because the SSL-Verification toolbar could not distinguish phishing sites from legitimate sites that do not use SSL. Sadly, many such sites exist in the wild, and some were used in our study. But even when the toolbar functioned properly, it was often ignored. One subject commented that the toolbar looked like an advertisement banner, so it was unclear whether it was put there by the browser or by the site.

The other two toolbars were thought more effective than the browser's own address bar. A common remark on the security toolbar was that the toolbar worked with the address bar: the toolbar alerted and warned the subject, causing the subject to pay more attention to the address bar.

Some subjects did not know how to interpret the information the toolbars displayed—especially the Neutral-Information toolbar. One subject said: “How do I have any idea about the [registration] time and location of a site?”

Why Don't the Security Toolbars Work?

Many users relied on the web content to decide if a site is authentic or phishing. The web content has a large display area and is in the center of the user's attention. It can make itself very convincing. Most of the time, the web appearance does reflect the site's identity because of the low phishing rate in the real world. What's more, in the early days of phishing, phishing attacks frequently had poor grammar and spelling mistakes. In our study, simulated phishing sites had high-fidelity content. As a result, even though the security toolbar and other security indicators in the browser tried to alert the user, many users disregarded the security toolbars because the content looked so good.

Poor web practices on the part of e-commerce firms make phishing attacks even more likely to succeed. For example, many legitimate companies do not use SSL to protect their

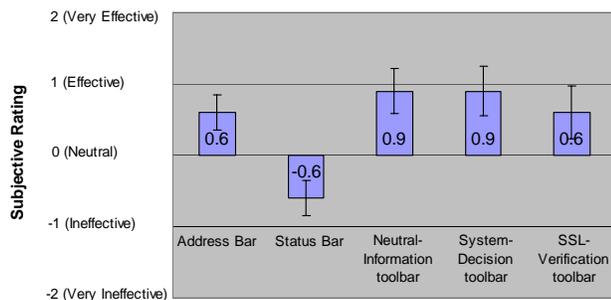


Figure 8. Subjective ratings of the address bar, the status bar and the toolbars

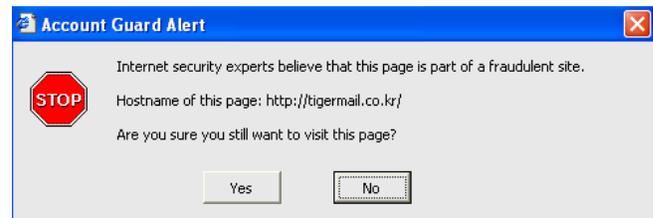


Figure 9. A sample blocking warning box at a phishing site

login page; this was a serious problem for the SSL-Verification toolbar. Many operators use domain names that are vague, inconsistent, or otherwise unrelated to their brands. Many organizations make their outsourcing relationships directly visible to Internet users. Such practices make it even harder for users to distinguish legitimate websites from malicious attacks.

FOLLOW-UP STUDY

A more effective interface for getting the user's attention about a phishing web site is to actually *block* access to it—for example, by popping up a modal dialog box when the site is visited. Several security toolbars, including Netcraft Toolbar, eBay Account Guard and SpoofGuard, display a pop-up warning when they have high confidence that the current site is phishing. This warning is likely to get the user's attention since it appears in the center of the browser and impedes progress until it is acknowledged.

A pop-up is a very aggressive warning, disrupting the user's task, so it must be accurate or it will be disabled. Since phishing attacks evolve rapidly, we have found that security toolbars are rarely certain enough about a new phishing attack to display a pop-up. As a result, these toolbars depend more heavily on the persistent toolbar display to warn users about new dangers. This is why our first study focused on the toolbar display.

Nevertheless one might expect a pop-up dialog to be more effective at prevent phishing attacks. To find out, we ran a follow-up study with new subjects to test the pop-up alert technique. The second study used the same scenario and the same attacks with the same numbering and positioning of attacks. Half of the subjects saw a blocking warning box (figure 9) at the phishing sites, which closely resembles the warning used by the Netcraft Toolbar. The rest acted as a control, using only a standard browser interface with the address and status bars are the main security indicators.

The follow-up study had 20 subjects aged 19 to 37 (average age 23). 18 (90%) were college students, 13 male.

It turned out that these subjects exhibited more caution than the subjects in the first study. We quantified their degree of caution as a sum of three indicator variables: 1 point if the subject ever tried to log out of a site; 1 point if the subject unchecked “remember me” at a login page; and 1 point if the subject failed to finish the task at a good site because of security concerns. The second study's subjects had an average caution score of 2.0, compared to 1.3 for the first

study, a significant difference (two-tail $t(39) = -3.29$, $p = 0.002$). Possible reasons for the difference include the demographics of the subjects and substantial media coverage about phishing and identify theft in the intervening three months.

As expected, the blocking warning box dramatically decreased the spoof rate of the wish-list attacks in the study, as shown in figure 10. The decrease was statistically significant (one-tail $t(9) = -2.88$, $p = 0.01$).

Seven out of the 10 subjects with the regular browser interface were spoofed by at least one wish-list attack:

- Six subjects (86%) said that the web content looked good or the same as they had seen before.
- Two subjects (29%) rationalized the suspicious URL. One subject, experiencing a similar-name attack at `www.walmart.com` by `www.walmart.com.global-update2.com`, said that “global-update2 is a service to do the website’s global updating and this service is working for Wal-mart.”
- Three subjects (43%) did not look at the URL in the address bar at all. One said that “I did not bother to look at the address bar since the page looked so good.”

Two subjects with the regular browser interface were spoofed by the PayPal attack, both PayPal users in real life. They mentioned that the site looked just like PayPal’s and that the phishing email was a reasonable request by PayPal.

The results from the 10 subjects who used the regular browser interface supported our conclusions from the first user study: many users depend on the web content to authenticate the site’s identity. Even though they are cautious and notice suspicious signs from the browser’s security indicators, since these signals are weak compared to the strong signals from convincing web content, the users tend to ignore or explain away the security indicators.

Of the 10 subjects who used the blocking warning box, none were spoofed by the PayPal attack but four were spoofed by wish-list attacks:

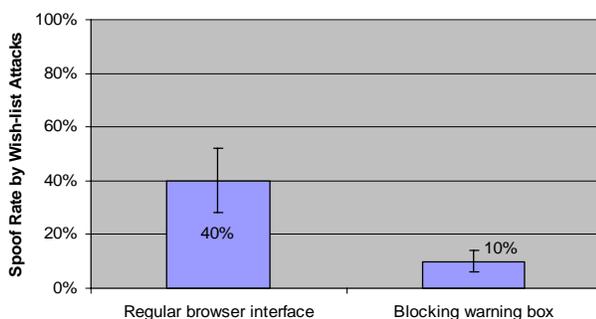


Figure 10. Spoof rates with a regular browser interface and the blocking warning box

- *None* of the four spoofed subjects offered that the content of the page was convincing as a reason that they were spoofed—somewhat ironic, since the content was in fact the real site! Apparently the warning box blocking the page is a stronger signal than the web content.
- Two subjects believed the warning and knew the site was phishing, but still wanted to complete the task. The subjects claimed that a wish list was not sensitive enough for them to decline John’s request. Apparently they did not realize that revealing John’s shopping password to an attacker could result in financial loss if John used the same username and password at another site.
- The other two subjects did not trust the blocking warning box. One said that he had never seen such a warning before. The other thought that the warning was wrong. This subject had his own anti-phishing strategy: he typed a wrong password at the suspicious site’s login page. If the site accepted the wrong password, he inferred that the site was phishing. But if the site rejected the wrong password, he concluded that the site was good and that the warning box was making an error. Clearly, this strategy does not work against phishing sites executing man-in-the-middle attacks, for these sites pass usernames and passwords on to the real site to perform the validity check. Interestingly, two other subjects in the follow-up study also used this same strategy to check a site’s authenticity; we never saw this behavior in the first study.

The follow-up study confirms that many users do not know how sophisticated a phishing attack can be. Some subjects were impressed by our simulated phishing attacks. One typical comment: “I cannot imagine that an attacker can make the attack so elegant to mirror a whole site.” Others wrongly believed that phishing sites cannot check password validity since they don’t have the correct password.

CONCLUSIONS AND RECOMMENDATIONS

We evaluated three types of security toolbars, as well as browser address and status bars, to test their effectiveness at preventing phishing attacks. All failed to prevent users from being spoofed by high-quality phishing attacks.

Users fail to continuously check the browser’s security indicators, since maintaining security is not the user’s primary goal. Although users sometimes noticed suspicious signs coming from the indicators, they either did not know how to interpret the signs or they explained them away. Many users had no idea how sophisticated an attack could be, and do not know good practices for staying safe online.

Design Principles for Anti-phishing Solutions

Based on these studies, we propose the following design guidelines to make effective anti-phishing solutions. We are also developing new techniques that use these guidelines.

As the follow-up study shows, active interruption like the popup warnings is far more effective than the passive warnings displayed in the toolbars. But it’s well-known that

popup confirmations, used indiscriminately, become less effective over time: the more often they appear, the less often users heed them. [17] In order to make the interruption effective, it should always appear at the right time with the right warning message. For example, most phishing attacks trick users into submitting their personal or financial information through web forms. Instead of using generic warnings like "Are you sure you want to continue sending this information over an unencrypted connection?", the browser should interrupt the user *only* for a dangerous action, like submitting one site's login information to another site. Knowing the user's intention makes it easier for the browser to protect the user.

User intentions should be respected. Simply warning users that something is wrong and advising them not to proceed is not the right approach. Users will take risks to finish the tasks they think worthwhile and are not good at evaluating the tradeoffs between the claimed benefits and the potential risks. Warnings that propose an alternative path (*e.g.*, directing users to the real intended site) allowing users to finish the task safely would probably be more effective.

If users must make security-critical decisions, it is best to integrate the security concerns into the critical path of their tasks so that they *have* to deal with it, and can't simply ignore it. Do not expect users to keep a separate security task in mind. Asking users to choose a safe mode to finish their tasks has been found to be more dependable and effective than merely reminding them to finish their tasks in a safe mode. [24]

Finally, Internet companies need to follow some standard practices to better distinguish their sites from malicious phishing attacks. Companies should use a single domain name that matches their brands name rather than using IP addresses or multiple domain names for servers. They should use SSL to encrypt every web page on their sites. SSL certificates should be valid and from widely used CAs.

Acknowledgements

We gratefully acknowledge the help and suggestions of Ruth Rosenholtz, Srinivas Devadas, the members of the UID group at CSAIL, and the anonymous reviewers. This work was supported in part by the National Science Foundation under grant IIS-0447800. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

1. Anti-Phishing Working Group. eBay – NOTICE eBay Obligatory Verifying - Invalid User Information. March 9, 2004. http://www.antiphishing.org/phishing_archive/eBay_03-09-04.htm
2. Anti-Phishing Working Group. Phishing Activity Trends Report, March 2005. http://antiphishing.org/APWG_Phishing_Activity_Report_March_2005.pdf
3. Bank, D. 'Spear Phishing' Tests Educate People About Online Scams. *The Wall Street Journal*. August 17, 2005.
4. BBC News. Passwords revealed by sweet deal. <http://news.bbc.co.uk/1/hi/technology/3639679.stm>
5. Chou, N., Ledesma, R., Teraguchi, Y., Mitchell, J.C. Client-Side Defense Against Web-Based Identity Theft. *11th Annual Network and Distributed System Security Symposium* (2004).
6. Dhamija, R. Tygar, J.D. The Battle Against Phishing: Dynamic Security Skins. *Symposium on Usable Privacy and Security* (2005), pp. 77-88.
7. eBay Toolbar and Account Guard. <http://pages.ebay.com/help/confidence/account-guard.html>
8. Emigh, A. Online Identity Theft: Phishing Technology, Chokepoints and Countermeasures. ITTC Report on Online Identity Theft Technology and Countermeasures. October 3, 2005. <http://www.antiphishing.org/Phishing-dhs-report.pdf>
9. Federal Bureau of Investigation, Department of Justice. FBI Says Web 'Spoofing' Scams are a Growing Problem. 2003. <http://www.fbi.gov/pressrel/pressrel03/spoofing072103.htm>
10. Fluendy, S. Phishing targeting online outlets. Computer Crime Research Center. March 16, 2005. <http://www.crime-research.org/news/03.16.2005/1050/>
11. Fogg, B.J, et al. What makes Web sites credible?: a report on a large quantitative study. *CHI* 2001, pp. 61-68.
12. Google Safe Browsing for Firefox. 2005. <http://www.google.com/tools/firefox/safebrowsing/>.
13. Herzberg, A., Gbara, A. TrustBar: Protecting (even Naïve) Web Users from Spoofing and Phishing Attacks. 2004. <http://www.cs.biu.ac.il/~herzbea/Papers/ecommerce/spoofing.htm>.
14. Jagatic, T., Johnson, N., Jakobsson, M., Menczer, F. Social Phishing. School of Informatics & Dept. of Computer Science, Indiana University. 2005. http://informatics.indiana.edu/fil/Net/social_phishing.pdf
15. Leyden, J. US phishing losses hit \$500m. *The Register*. September 29, 2004.
16. Netcraft Toolbar. 2004. <http://toolbar.netcraft.com/>.
17. Norman, D. A. Design rules based on analyses of human error. *CACM*, v26 n4 (April 1983), pp. 254-258.
18. PassMark. 2005. <http://www.passmarksecurity.com/>.
19. Sharif, T. Phishing Filter in IE7, September 9, 2006. <http://blogs.msdn.com/ie/archive/2005/09/09/463204.aspx>
20. SpooFStick. 2004. <http://www.spooFstick.com/>.
21. Sullivan, B. Consumers still falling for phish. MSNBC. July 28, 2004. <http://www.msnbc.msn.com/id/5519990/>
22. Whalen, T., Inkpen, K. Gathering Evidence: Use of Visual Security Cues in Web Browsing. *Graphics Interface 2005*.
23. Whitten, A., Tygar, J.D. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. *8th Usenix Security Symposium*, 1999, pp. 169-184.
24. Wu, M., Garfinkel, S., Miller, R. Secure Web Authentication with Mobile Phones. *DIMACS Workshop on Usable Privacy and Security Software*, 2004.