Effectiveness of a Phishing Warning in Field Settings

Weining Yang[1], Jing Chen[2], Aiping Xiong[2], Robert W. Proctor[2], Ninghui Li[1]
[1]Department of Computer Science
Purdue University
{yang469, ninghui}@cs.purdue.edu
[2]Department of Psychological Sciences
Purdue University
{chen548, xionga, rproctor}@purdue.edu

ABSTRACT

We have begun to investigate the effectiveness of a phishing warning Chrome extension in a field setting of everyday computer use. A preliminary experiment has been conducted in which participants installed and used the extension. They were required to fill out an online browsing behavior questionnaire by clicking on a survey link sent in a weekly email by us. Two phishing attacks were simulated during the study by directing participants to "fake" (phishing) survey sites we created. Almost all participants who saw the warnings on our fake sites input incorrect passwords, but follow-up interviews revealed that only one participant did so intentionally. A follow-up interview revealed that the warning failure was mainly due to the survey task being mandatory. Another finding of interest from the interview was that about 50% of the participants had never heard of phishing or did not understand its meaning.

Categories and Subject Descriptors

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

Keywords

Phishing; Phishing warning

1. INTRODUCTION

Phishing is a continuously growing and evolving threat in cyber security. It can be defined as: "Fraud perpetrated on the Internet; *spec.* the impersonation of reputable companies in order to induce individuals to reveal personal information, such as passwords and credit card numbers, online" [1]. Usually, phishing attacks start from email spoofing or instant messages, and the emails or messages contain links directing users to "fake" sites, where they are asked to provide personal information. As the fake sites look identical to legitimate ones, the users are tricked into entering sensitive information, which is stolen by the attackers.

HotSoS '15 , Apr 21-22, 2015, Urbana, IL, USA
ACM 978-1-4503-3376-4/15/04
http://dx.doi.org/10.1145/2746194.2746208.

Previous studies showed that users are easily deceived in phishing attacks, as suggested by a large reply rate for phishing emails, especially those from a friend's spoofed address [5]. It also has been found that users generally do not look at browser-based cues such as the address bar, status bar, and the security indicators [3]. Further, active phishing warnings, which force users to notice a warning by interrupting their current activity, showed a substantial compliance rate [4]. However, the same study also reported that users usually do not understand the meaning of the warnings.

Generally speaking, methods detecting phishing websites fall into two categories: blacklist-based methods and heuristic methods. The former leverage human-verified phishing URLs, which reduces false positive rate significantly [6]. However, the blacklist-based methods fail to detect newly created phishing sites. The heuristic methods identify phishing pages by analyzing page content through taking advantage of machine learning algorithms [2]. Heuristic methods are generally heavy-weight and thus, cannot be deployed on the client end. Also, such detection is relatively difficult for typical users to understand.

In this work, we proposed a light-weight client-end active warning, which can be easily deployed in a browser extension to protect users from phishing attacks. We examined the efficacy of the mechanism in a real-world field setting. We recorded participants' responses to two simulated phishing attacks during a 6-week period and interviewed them at the end of the study.

2. BROWSER EXTENSION DESIGN

Presumably, phishing sites are infrequently visited and the targeted websites of phishing attacks are popular ones. Our phishing detection is based on the difference of popularity of phishing sites and legitimate ones. We crawled phishing URLs in Phishtank.com from Sep 21 to 30, 2014, and obtained 20,797 unique URLs in total. For each phishing URL, we examined its domain on Alexa.com and got its rank, which was the smaller one of its global rank and region rank. The distribution of rank is listed in Table 1. More than 91% of phishing sites were set up in domains with rank greater than 10,000. Also, we manually examined the reported URLs from popular sites, and most of the pages were removed within one hour. Therefore, we chose to warn user about domains with ranks greater than 10,000.

A Chrome extension was developed to present warnings to participants. If one visits an unpopular site, i.e., the rank of its domain is larger than 10,000, and attempts to enter any information on the page, a warning pops up. The warning is

Table 1: Rank distribution of phishing domains

Rank	Frequency
1-100	510(2.5%)
101-1000	353(1.7%)
1001-10000	899(4.3%)
10000-100000	918(4.4%)
100000-1000000	699(3.4%)
1000000+	17418(83.8%)

Figure 1: The warning display, depicted for a domain called xorbin.com

illustrated in Figure 1. Because ranking itself is too technical for typical users, the warning does not explicitly display the site's ranking. Instead, the domain name is extracted from the URL and marked in different colors in order to aid participants to determine the current webpage's legitimacy. Participants can get a detailed description about phishing by clicking the "Why this warning" button. Also, the detailed ranking information will be displayed if the "About this domain" button is clicked. Participants have three options to choose when they see the warning. The first option adds the domain in a whitelist so that the warning will not pop up on the same domain again. The second option allows participants to visit the page only for this time. When participants try to visit the same domain again, the warning will also be shown if they attempt to key in any information. The third option is simply to close the webpage.

3. STUDY DESIGN

We designed a study simulating phishing attacks in real-world field settings. The study lasted 6 weeks. Participants were recruited by way of fliers placed around the campus. At the beginning of the study, participants were required to come to our lab to install the Chrome extension. Participants were told that the study was mainly regarding Internet browsing behaviors. At this point, nothing about phishing was mentioned. They were asked to register new accounts on our survey website. An email address and a password were required in registration. Every week, participants were required to login to our website and finish a short questionnaire, which asked for an estimation of time spent on different events on the Internet. The link of the questionnaire was sent via email. Normally, the link in the email directed them to a website under the domain of "purdue.edu". In weeks 4 and 6, the links in the email were associated with two newly registered "phishing" domains maintained by us. We recorded all the actions on our phishing sites.

Participants were divided into two groups: control group, in which no warning was presented during the whole experiment time; experimental group, in which participants saw the warning described in section 2 when they attempted to enter information on domains ranked greater than 10,000, as well as on the "phishing" domains we created. At the end of the study, participants came to our lab again for a semi-structured interview, after which they were debriefed about the true purpose of the study.

4. PRELIMINARY RESULTS

In this pilot study, we recruited 9 participants, 3 of whom were in the control group. During the "phishing" weeks, 2 participants in the control group provided their passwords directly while only 1 participant in the experimental group typed in the correct password in the first attempt. However, for the participants who saw warnings, 5 of them chose to permanently trust our "phishing" domain; the rest selected the second option and continued entering information. No one chose "Close the page" on the warning interface or closed the tab in the browser.

To further understand the rationale behind the participants' actions, we interviewed them regarding phishing and the warning generated by our Chrome extension. It turned out that only one participant intentionally provided a wrong password, while the rest entered mismatched passwords by accident. Most of the participants said they did not know the meaning of our warning and tended to ignore it in part due to the interface design. Another finding was that about half of the participants indicated that they had not heard about or did not know the meaning of phishing.

5. NEXT STEP

The results demonstrated that the task, the warning interface, and the knowledge of phishing are critical factors that should be taken into account during a simulated phishing study. Consequently, we have redesigned the method for a full study that we are currently conducting. This study is using a redesigned warning interface and a different scenario, in which the phishing message replicates a popular commercial website promotion requesting a voluntary response.

6. REFERENCES

[1] *The Oxford English Dictionary*. Oxford University Press, 2015.

[2] A. Bergholz, J. H. Chang, G. Paaß, F. Reichartz, and S. Strobel. Improved phishing detection using model-based features. In *CEAS*, 2008.

[3] R. Dhamija, J. D. Tygar, and M. Hearst. Why phishing works. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 581–590. ACM, 2006.

[4] S. Egelman, L. F. Cranor, and J. Hong. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1065–1074. ACM, 2008.

[5] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer. Social phishing. *Communications of the ACM*, 50(10):94–100, 2007.

[6] S. Sheng, B. Wardman, G. Warner, L. Cranor, J. Hong, and C. Zhang. An empirical analysis of phishing blacklists. In *Sixth Conference on Email and Anti-Spam (CEAS)*. California, USA, 2009.