Testing PhishGuru in the Real World

Ponnurangam Kumaraguru, Steve Sheng, Alessandro Acquisti,
Lorrie Faith Cranor, Jason Hong
Carnegie Mellon University
ponguru@cs.cmu.edu, shengx@cmu.edu, acquisti@andrew.cmu.edu,
lorrie@cs.cmu.edu, jasonh@cs.cmu.edu

ABSTRACT

In real world testing of PhishGuru, an embedded training system that teaches people how to protect themselves from phishing attacks, we found (a) PhishGuru is effective in training people in the real world; (b) users retained knowledge when trained with PhishGuru in the real world; (c) a large percentage of people who clicked on links in simulated emails proceeded to give some form of personal information in the real world; (d) people trained with spear phishing specific training material did not make better decisions in identifying spear phishing emails compared to people trained with generic training material. We also observed that PhishGuru training could be effective in training other people in the organization who did not receive training messages directly from the system.

1. INTRODUCTION

PhishGuru is an embedded training system in which users are sent simulated phishing attacks and are presented with training materials when they fall for the attacks. Prior laboratory studies have shown that training users with PhishGuru is an effective way to teach users to make better decisions in identifying phishing scams. PhishGuru motivates users to pay attention to training materials by taking advantage of teachable moments. Our goal in this research is to demonstrate the effectiveness of PhishGuru training in field trials. This study is the first empirical investigation of PhishGuru in the real world.

To evaluate PhishGuru in the real world, we conducted a study among employees in a Portuguese company. The simulated phishing emails we used in this study consisted of only spear phishing emails targeted at the employees of the company. To investigate the effect of different training messages, we used one that had instructions on how to protect against regular phishing scams (generic training) and one that had instructions for protecting against spear phishing scams (spear training).

2. STUDY SETUP

This study was conducted at a Portuguese company. All training materials and simulated phishing emails were translated into Portuguese. The study included three conditions: “control,” “generic training,” and “spear training” Participants in the control condition did not receive any training. Participants in the generic training condition received a simulated spear phishing email and saw the training material with generic instructions when they clicked on a link in the email. Participants in the spear training condition received a simulated spear phishing email and saw the training material with specific instructions pertaining to spear phishing emails when they clicked on a link in the email. Over the next 10 days we sent participants 3 additional emails to test whether the training was effective as shown in Table 1. We gave all employees the option of filling out the exit survey on day 20. We excluded the employees from the technical department or employees with technical background such as in computer science or engineering for any analysis that is discussed in this paper. We found that employees in the technical department or employees having technical background were able to identify the phishing emails correctly that were part of the study. From here on all data presented in the paper is only from departments other than technical. We had 67 participants in the control condition, 64 in the generic, and 65 in the spear training condition.

3. Hypotheses

We tested the following three hypotheses in this study:

Hypothesis 1*: A large percentage of people who click on links within simulated emails proceed to give some form of personal information in the real world.*

Hypothesis 2*: PhishGuru (embedded training) is effective in training people in the real world.*

Hypothesis 3*: People trained with spear training materials make*

Table 1: The type of the emails that were used in the study, days in which the email was sent and the conditions to which the emails were sent.

Emails	Type	Day of sending	Conditions	Relevant features of the email
Train	Spear phishing	Day 0	Generic and spear training	Asked to enter their user name and password in order to use the corporate network
Test-1	Spear phishing	Day 2	All	Internal network password expired; asked to change their password
Test-2	Spear phishing	Day 7	All	Asked to update their communication information
Test-3	Legitimate-with-link	Day 10	All	Asked to read the company’s updated security policy

Table 2: Mean correctness for phishing emails; values presented in percentage. There was significant difference between Day 0 and Day 2 in both generic and spear conditions.

	Day 0	Day 2	Day 7
Control	X	70.15 (47)	80.60 (53)
Generic training	57.81 (37)	81.25 (52)	79.69 (51)
Spear training	64.62 (42)	87.69 (57)	83.08 (54)

Table 3: Percentage of trained employees who clicked on the testing emails; values in the brackets are the number of people. There was no significant difference between the conditions on all three emails.

Training conditions	% clicked on the training email	% of trained users who clicked on Day 2 email	% of trained users who clicked on Day 7 email
Generic	42.19 (27)	18.51 (5)	14.2 (6)
Spear	35.38 (23)	21.73 (5)	11.4 (4)

better decisions in identifying spear phishing emails compared to people trained with generic training materials.

4. RESULTS

In this study we found that a large percentage of the participants who clicked on links in emails went ahead and gave some form of personal information to the phishing websites. As we had access to the information that was entered into phishing websites, we were able to check the usernames and other details that were entered. We found around 80% of the participants gave some form of personal information to the phishing websites. In past work, researchers have shown this percentage to be 93 and 90 in laboratory studies [1], [2]. This confirms Hypothesis 1. Even if users do not provide personal information, they may still place themselves at risk of malware if they click on a phishing link. For the rest of this paper we consider someone to have fallen for a phishing attack if they click on a link in a phishing email, regardless of whether they go on to provide personal information.

In both the training conditions (generic and spear), participants acquired and retained knowledge after 7 days of training. We found significant difference in the mean correctness between Day 0 and Day 2 in both the training conditions. Mean correctness is a correct decision that participants make on their emails. For example, deciding that a phishing email is a phishing email and therefore deleting the email without clicking on a link. Table 2 presents the mean correctness for all three conditions on different days of the study. This result supports the hypothesis that participants in the generic and spear training conditions were able to make improved decisions immediately after being trained and they were able to retain the knowledge after 7 days. Table 3 presents the percentage of employees who clicked on the training email and percentage of those trained employees who clicked on the testing emails (Day 2 & Day 7). We also found no significant difference between the training conditions (generic and spear) in the ability to identify phishing emails after the training. These results lend support to Hypothesis 2 and reject Hypotheses 3.

Only three employees among all the three conditions clicked on the legitimate-with-link email that was sent on the Day 10. To test this behavior, we sent a different legitimate-with-link email (asking to read the sales report of the company) on Day 14. Again we found only 3 employees clicked on the link in this email. Unfortunately, we don't know how many employees clicked on links in legitimate company email prior to training, but, we suspect that most of them do not.

We observed that studying phishing training is difficult in the real world due to employees discussing the training messages among themselves. But, this suggests that PhishGuru training can be effective in training people who are not part of the study. This may be the reason for high mean correctness in the control group on Day 2 and Day 7 (Table 2). None of the employees completed the exit survey.

5. RECOMMENDATIONS

This study was the first empirical investigation of PhishGuru in the real world. There were many lessons that we learned from this study. The following are a few recommendations for conducting real world embedded training studies.

- *Content of the email is important*: The simulated emails that are used in the study should be relevant and have a compelling argument for participants to make a decision.
- *Incentive for participants*: The employees who are part of the study may not have the incentive to provide feedback or complete an exit survey. So, providing some form of incentive (cash or prize) to the participants is necessary.
- *Avoid experts*: Technically savvy people (experts) don't click on links in emails, so recruiting them as participants for phishing (embedded training) studies should be avoided. This may reduce the data that one has to throw away.
- *Use participants from different locations*: Although it is difficult in a real world setting, it will be useful if participants of the study are chosen from different locations. So they don't talk to each other about the study.
- *Keep it simple*: The companies that agree to real world studies may not have incentives to collect data at the level of detail that researchers would want. Therefore, the procedures for collecting data should be minimized and made simple.

6. ACKNOWLEDGMENTS

Our thanks to Members of Supporting Trust Decisions for fruitful discussions on this topic.

7. REFERENCES

[1] Kumaraguru, P., Y. Rhee, A. Acquisti, L. Cranor, J. Hong, and E. Nunge. 2007. In *Proceedings of CHI 2007.* Protecting People from Phishing: The Design and Evaluation of an Embedded Training Email System.

[2] Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L. F., and Hong, J. 2007. Getting users to pay attention to anti-phishing education: Evaluation of retention and transfer. e-Crime Researchers Summit, Anti-Phishing Working Group.