1-1-2009

The Effectiveness of Deceptive Tactics in Phishing

Kent Marett  
Mississippi State University, kmarett@cobilan.msstate.edu

Ryan Wright  
University of San Francisco, rwright@usfca.edu

Follow this and additional works at: http://aisel.aisnet.org/amcis2009

Recommended Citation
http://aisel.aisnet.org/amcis2009/340

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2009 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
The Effectiveness of Deceptive Tactics in Phishing

Kent Marett  
Mississippi State University  
KMarett@cobilan.msstate.edu

Ryan Wright  
University of San Francisco  
rwright3@usfca.edu

ABSTRACT

Phishing, or the attempt of criminals to obtain sensitive information through a variety of techniques, is still a serious problem for IT managers and Internet consumers. With over 57 million Americans exposed to phishing in 2005, a reported 5% of recipients were victimized. Some believe that one percent of all email is phishing-related, and estimates of financial losses vary from 100 million to 1 billion dollars (US) a year (Goth, 2005). Our research examines the properties in a phishing email that may or may not influence the users to give out personal and sensitive information. For this field experiment we use students to test the effect that certain types of content have on the phishing process. The study outcomes suggest that user’s do not pay attention to the sender’s domain in a phishing email but do respond to personalized messages and messages that demand an immediate response.

INTRODUCTION

For several years now, both computer industry sources and the popular press have informed the computing community about intentionally deceptive practices on the Internet. These practices include technologically catchy names such as phishing, pharming, page-jacking, mouse-trapping, etc. The basic premise of these deceptive techniques is for the person behind the fraudulent effort to masquerade as a credible figure in order to acquire sensitive information. This technique is based on social engineering whereby sensitive information, including credit card numbers, passwords or social security numbers, can be elicited from unsuspecting persons. While the technique of social engineering is not new, the specific contextualization of the Internet has changed the conditions under which it can occur. No longer is face-to-face or telephone interaction the media of choice for social engineers; it is now the web or email.

In order to fully explain the social engineering aspects of deceptions that occur on the Internet this manuscript focuses on the act of “phishing.” Phishing, a term whose origin arises from the sophisticated technical lures to “fish” for a users’ sensitive information, is defined as a practice of directing users to fraudulent web sites for the purpose of obtaining personal information (RSA Security, 2004). The main technique used by phishers is conducted via email and is known as the bait and hook method. In this method, users receive an unsolicited email from an agent posing as a legitimate party, which is the baiting aspect of phishing. Typically, the message either directly asks for information or directs the user to a web site that is designed to collect the information, which is considered the hook portion of phishing. With over 57 million Americans exposed to phishing in 2005, a reported 5% of recipients were victimized. Some believe that one percent of all email is phishing-related, and estimates of financial losses vary from 100 million to 1 billion dollars (US) a year (Goth, 2005). The consulting company Gartner is even more pessimistic, estimating that 3.6 million adults combined to lose more than 3 billion dollars to phishing attacks in the 2007 calendar year (Gartner, 2007).

The technical literature on Internet security is increasing, though little is known about the behavior and cognitive psychological aspects of the Internet user that are exploited by deceptive efforts like phishing. Typically past research has sought to understand the HCI properties of the phishing web site. Past evidence suggests that phishing web sites and actual web sites are very hard to distinguish by typical users (Dhamija & Hearst, 2006), and even sophisticated automated systems designed for detecting bogus websites (Liu, Deng, Huang, & Fu, 2006) and anti-phishing toolbars made specifically for determining the legitimacy of websites (Wu, Miller, & Garfinkel, 2006) are inconsistent for warning users of problems. It seems that, once Internet users arrive at the fraudulent site, there is very little chance for them to detect the true duplicitous nature of the request. There may be a much better chance for detecting the false nature of the message by examining the original email communication. Accordingly, we recently conducted an experimental study designed to investigate why experienced internet users act on deceptive bait and hook email messages, focusing particularly on the “bait” portion of the message. We were interested in the deceptive content contained within the message and, more specifically, what content is the most effective and what content is most detectable. Many phishing messages, such as the one displayed in Figure 1
below, are plagued with poor grammar, which effectively cancels any credibility established with the use of stolen logos and wordmarks. An observant receiver of this email will become suspicious upon reading such atrocious language, but assuming phishers start using proper grammar in the future, are there any other content-based indicators that might be noticeable to phishing recipients?

The goal of this study is to examine the indicators that result from the tactics used to craft a deceptive message (like that found in a phishing effort) for reliable cues that message receivers can use for deception detection. The paper is organized as follows. We begin with a literature review that ties phishing to research on deceptive tactics, specifically a taxonomy of tactics originally developed by Johnson, Grazioli, Jamal, and Berryman (2001). That is followed by two hypotheses formulated to explore the relationships between those tactics and deception detection. The method used to test those hypotheses and the findings that resulted follow, concluding with a discussion section which includes our thoughts on future research on this topic.

![Figure 1. Actual Phishing Message Received by the Authors.](image)

**PHISHING TACTICS**

A taxonomy developed by Johnson and colleagues (2001) provides some of the more common methods of manipulating content so as to appear legitimate to unsuspecting readers. Grounded in their Theory of Deception, the taxonomy identifies tactics which deceivers use to either prevent message receivers from understanding the full nature of the item being discussed (deception by omission) or to give the message receiver a false representation of the item (deception by commission) (Johnson et al., 2001; Grazioli & Jarvenpaa, 2003). As Burgoon et al. (2008) suggest, deceptive tactics like those found in the taxonomy are effective because they seek to avoid deviating from normative behavior, a strategy which follows from expectancy violations theory. By using these tactics to craft a deceptive, yet outwardly realistic, message, deceivers aim to prevent triggering a state of suspicion in the message receiver (Zhou, Burgoon, & Twitchell, 2003).
The Johnson et al. taxonomy, displayed in Table 1 below, describes seven different deceptive tactics that were originally conceived and observed in fraud accounting, and have themselves been used in experiments involving the falsification of personnel information within databases (Biros, George, & Zmud, 2002) and on fraudulent eCommerce websites (Grazioli et al. 2003; Riquelme & Kegeng, 2004). As we describe in this section, the seven tactics are also applicable to phishing efforts, but four of the seven tactics are relevant to a specific type of content contained within a phishing message: decoying, inventing, dazzling, and mimicking.

It is unclear whether or not phishers purposely employ content-based tactics in their attempts, but they are certainly present at times, such as in the example in Figure 1. The dazzling tactic can be seen in many phishing messages, particularly with the mention of items or properties that the target should be familiar with. The simple mention of a well-known name or event can give the air of “inside knowledge” that many social engineers strive to convey. The use of the decoying tactic, such as personalizing a message for a specific recipient, is often used to distract the victim’s attention away from the false content of the message. By addressing the message using the recipient’s name or by including other personal information, like a phone number, phishers hope to both distract the reader from deceptive information and add a measure of authenticity. Mimicking a sender’s address is a part of every phishing attack, as the assumption of a trusted person or institution’s identity is necessary to blur the true source of the message, as well as to give the message the needed veracity that is highly successful in phishing. Finally, an inventing tactic, such as the introduction of an urgent deadline for a response, is also commonly seen in phishing, as it purposely presents a clear fabrication about the topic of the message (Jagatic, Johnson, Jakobsson, & Menczer, forthcoming).

<table>
<thead>
<tr>
<th>Deception Tactic</th>
<th>Definition</th>
<th>Phishing Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Content-related</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dazzling</td>
<td>Obscuring or blurring information about the deception core by adding information, not eliminating it.</td>
<td>“Name Dropping” – including a name that is known to the receiver. (e.g. a colleague)</td>
</tr>
<tr>
<td>Decoying</td>
<td>Distracting the victim’s attention away from what is really going on.</td>
<td>“Personalization” – adding the name and other personal information to the phishing email.</td>
</tr>
<tr>
<td>Mimicking</td>
<td>Assuming someone’s identity or modifying the core so it copies the features of a legitimate item.</td>
<td>“Sender address” – assuming a false name and email address of the message sender</td>
</tr>
<tr>
<td>Inventing</td>
<td>“Making up” information about the core. The core might not exist, or its characteristics might be utterly unrealistic.</td>
<td>“Call to action” – Introducing time pressure (e.g. a response is needed within 24 hours)</td>
</tr>
<tr>
<td><strong>Contextual</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masking</td>
<td>Eliminating or erasing crucial information so that representation of key aspects of the item does not occur, or produces an incorrect result</td>
<td>Typically used when failing to disclose important details, such as the true identity of the sender</td>
</tr>
<tr>
<td>Relabeling</td>
<td>Describing the situational context and its characteristics expressly to mislead.</td>
<td>Typically used to misrepresent the fictitious situation as being true</td>
</tr>
<tr>
<td>Double Play</td>
<td>Convincing the victim that he or she has been taken advantage of or that an unintentional mistake has occurred.</td>
<td>Typically used to motivate the recipient (social engineering)</td>
</tr>
</tbody>
</table>

Table 1. Adaptation of the Johnson et al. Taxonomy of Deceptive Tactics for Phishing.
The remaining three deceptive tactics in the Johnson et al. (2001) taxonomy, masking, relabeling, and double play, are also present in phishing efforts, but we consider them more contextual than content-based. Masking tactics do not involve manipulating content but instead involve the omission of information, and these contextual tactics are facilitated by the anonymity available on the Internet (Grazioli & Jarvenpaa, 2003a). With masking, the phisher omits valid information, such as his or her own identity or location, which is typical of all phishing efforts. Relabeling is just as common in phishing, in that the phisher attempts to make the situation being described as being true and accurate. Double play tactics are typical of the motivation used by social engineers, in that they induce action by persuading the recipient that acting is in his or her best interest (Mitnick, 2002; Grazioli & Jarvenpaa, 2003a). We believe that these three tactics must be used in every phishing effort and, thus, not content able to be manipulated by the phisher, which is the focus of this study.

RESEARCH MODEL

While the four content-based tactics are used with the intent of successfully deceiving others, it is possible that certain tactics can also lead to the downfall of the phishing effort. In other words, using tactics to make the phishing effort more believable may also make the deception more detectable, a conundrum that has been referred to as “the deceiver’s curse” (Grazioli & Carrell, 2002). According to Johnson and colleagues (2001), the detection of deceptive information, such as that found on “page-jacked” websites and phishing messages, is contingent on a sort of violation of expectancies. A piece of text-based information that has been manipulated using one of the above tactics can be defined as a “deception cue” (Grazioli, 2004). The deception cues are available to everyone who reads the information, and should the reader notice that the cues are inconsistent with the expected behavior or policies he or she has become accustomed to, the reader will likely be aroused to suspicion and investigate further. Unfortunately, not all message recipients notice the deception cues. Given the availability of the four content-based deceptive tactics for phishing efforts, we established the following two exploratory hypotheses:

H1: Content-based deceptive tactics increase the likelihood of successfully deceiving others with a phishing effort.

Further, we wished to compare the effectiveness of the individual tactics. To our knowledge, no prior research has investigated the content-based tactics in a phishing environment; therefore, we composed the following null hypotheses:

H2: Content-based deceptive tactics are equally effective for successfully deceiving others with a phishing effort.

In order to fully understand the effectiveness of these tactics, we tested them with an experiment involving subjects possessing information that “others” were interested in.

METHOD

Our intended targets were a recent class of 224 business students taking an introductory MIS course. The average age of the students was 21.02 years, and slightly more than half of them were male (52%). This class had a cross section of all majors in the business school. As a part of the requirements for the course, students completed assignments and took exams and quizzes using specialized software installed in the lab and accessible only through a password system. Unique passwords, which were referred to as the “Super Secure Code” or the “SSC,” were issued to each student the first week in class. We made clear the importance of keeping the SSC private. The SSC was sealed in an official university envelope that included the university logo and contact information for the IT department, and there was also a signature across the seal of the envelope. On the outside of the envelope was printed the subject’s name and student number. The code was printed on official university letterhead with the disclaimer “Do not disclose this code to anyone.” Finally, lab instructors explained to the students that they should not share this code with anyone under any circumstance, as it could “breach the secure grading process and could affect their grade for the class.” Students signed a non-disclosure agreement that they would abide by the class policy.

Eight weeks into the course, after class lectures on information security and online risks, students received an email from a fictitious IT employee, “Jason Roth,” requesting them to disclose their “super-secure code” in order to help recover from a data management accident. We utilized four tactics used by phishers in order to determine which tactic(s) were more effective in soliciting the “super-secure code”. The message originated from one of three email domains in an effort to mimic a sender’s information: a generic “mail.com” account, a spoofed address that appeared to be from the university.edu domain, and a legitimate.edu address coordinated with the university behind the scenes. The legitimate.edu phishing email was sent using the university’s mail servers. Spoofing, on the other hand, was performed by configuring an off-campus mail server to use an auto mail to send out emails that are addressed from an.edu user but are really not a part of the.edu domain.
The message content itself was manipulated using one of three other deception tactics. Figure 2 displays a sample message featuring all four of the content manipulations, mimicking the manipulated sender information, inventing (the urgent call to action vs. no urgent call), dazzling (name-dropping the course instructor’s name vs. no name-dropping), and decoying (personalization of the message vs. “Dear Student”). For comparison purposes, we randomly sent some students a message containing all four manipulations and yet other students a message that contained no manipulation other than the mimicked accounts. Having access to the accounts used to send the messages, we waited seven days to see which message was the most effective at eliciting the SSC and which message was the most likely to be detected as fraud. We considered the phishing effort a success when a student replied with the SSC as directed. A 4 X 3 factorial field experiment was set up. Where the mimicked was manipulated by the sender’s medium being: 1) a real university.edu account that was sent from a university server (real.EDU), 2) a spoofed university.EDU account that was sent from a spammer email server where any domain and email account could have been used as the sender (spoofed.EDU) and from a free email service mail.com (Mail.com). The four conditions for the content were: 1) personalized (e.g. addressed the subject directly in the email), 2) name dropped (e.g. used their instructor’s name and their classes rubric in the email), 3) urgency (e.g. asked the email be responded to immediately), 4) low condition that had no content other than the call to respond to the email. See figure 2 for examples of the conditions in an actual email.

![Figure 2. Example of the Phishing E-Mail.](image)

**RESULTS**

We have reported preliminary results in Table 2 below for each of the conditions and mediums and whether subjected responded with their SSC. We can see from this frequency table that it is unlikely that the medium (e.g. Real.EDU, spoofed.EDU and Mail.com) has any impact on whether subjects respond to the phishing email with their SSC. Further, there does seem to be differences in the effects of the type of content though. SPSS 16.0 (Green & Salkind 2005) was used to analyze the data.
Table 2. The Deceived (out of about 20 receivers per cell).

To test the hypotheses, we used a logistic model as the dependent and the independent variables were categorical (Pedhazur and Schmelkin 1991). Specifically, the independent variables represent the Medium Sent and the Content conditions, whereas the dependent variable was binary (answered with SSC = 1 and did not give out SSC = 1). For the purposes of this study the medium or mimic condition is defined as the server used to send the email (real.EDU, spoofed.EDU or Mail.com). The content type consisted of one of four treatments: a low (baseline) condition, a message personalized to the user (decoying), a message in which a familiar name was dropped (dazzling), and a message requesting an urgent reply (inventing). The Enter Wald method (Green & Salkind 2005) was used in the regression. This involves testing the model with all of the particular conditions. Table 3 shows the variables in the final equation. As the frequency statistics suggest, not all of the Medium conditions were a statistically significant factor in phishing deception. Further, name dropping was not a statistically significant variable in the regression equation. The final model was significant at p = .000 and had an R-squared of .263 (Mertler & Vannatta 2001).

Table 3. Logistic Regression Results for the Variables

Hypothesis 1 was partially supported as the content based approaches, other than name dropping, were far more successful than the low condition. The results suggest that the messages that name dropping had no statistic effect on the outcome. Each of the other content techniques were statistically significant factors in the logistic model. Often in regression analysis, other models are run that exclude the non-significant factors. This type of backwards modeling was not appropriate for this
study. It is impossible for us to exclude the medium conditions in the model as every email message sent in our study necessarily had a sending address.

Interestingly, the mimicking tactic made little difference, as students replied to “Jason Roth” fairly evenly across all three mail domains. The results suggest that inventing and decoying seemed to be the most productive tactics in this particular effort. This corresponds favorably to findings by Grazioli and Jarvenpaa (2003b), who found that inventing was one of the most frequently seen deception tactics online, as the fabrication of an event or circumstances can easily be applied to information intended for a mass audience, like the information found on a web page or in a phishing message. On the other hand, decoying is less often seen online, as it requires a deeper understanding of the target. In this case, decoying was effective, but the tactic required us to tailor each personalized message for specific targets, which is not feasible for a phisher attacking large numbers of email users to attempt. This could mean that, should phishers find an economical way to personalize messages in the future, the possibility for deception could become even greater than it currently stands. Overall, Hypothesis 2 was not supported, as some manipulations were more effective for deceiving recipients than others (see Table 4 below).

<table>
<thead>
<tr>
<th>Deception Tactic</th>
<th>Outcome</th>
<th>Standardized beta weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dazzling</td>
<td>It was not more successful at eliciting the “SSC” than the low condition.</td>
<td>Not Significant*</td>
</tr>
<tr>
<td>Decoying</td>
<td>Using the subjects named gained attention to the message.</td>
<td>-2.51</td>
</tr>
<tr>
<td>Mimicking</td>
<td>This condition surprise did not make a statistical difference.</td>
<td>Not Significant*</td>
</tr>
<tr>
<td>Inventing</td>
<td>The urgency of the message seemed to create an opportunity for phishers.</td>
<td>-1.87</td>
</tr>
</tbody>
</table>

*p<0.01

Table 4. Summary of the Outcomes.

Following debriefing, we required that the students locate the phishing message in their accounts, if possible. Using ANOVA we found that regardless of which type of mail server (hotmail, gmail, commercial, etc.) students used to collect incoming email, providing evidence there was no systematic difference between the mail servers and the emails that were unrecoverable (p=.89). In other words, there was no one email service that more ably filtered or rejected our phishing message than the other email services.

DISCUSSION

Our analysis of the deceptive cues presumes that phishers use well-planned, sophisticated methods for creating their messages, but in real life, that is anything but the case. As Berghel (2006) effectively demonstrates, the content of many phishing messages appears unprofessional, inappropriate for the situation, and grammatically horrific. The fact that actual phishing attempts that would, upon close inspection, make grade-school English students blush, yet successfully reel in unsuspecting recipients indicates that we should not expect better detection rates any time soon.

Although these results do suggest that techniques like personalization and applying pressure for a quick response seem to be successful for phishers, we believe that equal attention should be given to the mimicking tactic. Despite the fact that no one sender address was more effective than the others in our phishing effort, this could be a byproduct of the fact that the e-mail address was manipulated rather than the sender’s name, which can be just as easily mimicked. Often when viewing emails we do not see the domain, rather merely focusing upon the sender’s name. However, just as examining an embedded URL within the message can reveal the false nature of a message (Berghel, 2006). We believe that there is promising potential for detection if recipients better scrutinize the sender’s address to determine if it is valid.

From a phishing perspective, our attack was a complete success. We gained the super-secure codes of a quarter of the students in the course, as well as other unsolicited information, including some students’ cell phone numbers, student ID
numbers and even a couple social security numbers. This result is even more discouraging from an information security perspective, considering the students had signed non-disclosure agreement forms and were party to security and privacy lectures in the course. There was also an unexpected event that should have also heightened students’ awareness to phishing efforts. Two weeks before our attack, a large portion of students, faculty, and staff were subjected to the phishing message displayed in Figure 1. This was soon followed by email warnings about the attack sent to everyone with a university email account by the department of technology. Even with this event fresh in their minds, many students proceeded to disregard everything in order to help the well-meaning “Jason Roth.” While the prevailing logic seems to be that user education, backed by anti-phishing technology such as toolbars, are among the best countermeasures for preventing phishing success, we wonder if that is enough. Perhaps a training exercise like the one we performed can leave a more indelible lesson. We would like to think that our students would be less likely to fall for a phishing attack in the future. It is our hope that, by better understanding these tactics, IT managers can better tailor training to help combat phishing.

The results presented here offer avenues for future research. The deceptive tactics used here were followed the methods of social engineering used by phishers. The content of different phishing messages may change with time and opportunity, but it would be beneficial to identify social engineering tactics that better facilitate detection over time. One potential method might include developing reliable decision trees and models for handling text-based deception (Zhou et al. 2004). This might involve analyzing a corpus of phishing messages to determine the most commonly used tactics. Another potential avenue for future research in the possibility that specific mediums may interact with other content based phishing techniques. Finally, we should acknowledge the possibility that the behaviors and perceptions of undergraduate students could differ from those of more experienced computer users, so future research should also attend to generalizing the results found here to employees in organizational settings, and among people of different age groups and educational backgrounds.

REFERENCES


