Managing Phishing Emails: A Scenario-Based Experiment

M. Pattinson[1], C. Jerram[1], K. Parsons[2], A. McCormac[2] and M. Butavicius[2],
[1] Business School, University of Adelaide, Australia
[2] Defence Science and Technology Organisation, Edinburgh, Australia
e-mail: {malcolm.pattinson|cate.jerram}@adelaide.edu.au,
{kathryn.parsons|agata.mccormac|marcus.butavicius}@dsto.defence.gov.au

Abstract

In this paper, the authors report on a collaborative research project that investigates how people respond to phishing emails compared to genuine emails and what factors contribute to this behaviour. A scenario-based, role-play experiment was conducted by administering a web-based questionnaire via a series of seminars facilitated by a member of the research team. This questionnaire asked each of the 117 participants to evaluate 50 emails – half that were genuine and half that were phishing emails. In addition, demographic, personality and cognitive information was elicited from each participant. The major findings of this preliminary investigation indicate that personal characteristics such as familiarity with computers, extraversion, openness and cognitive impulsivity each have a significant effect on an email user's response to phishing emails.

Keywords

Information security (InfoSec), Information risk, Phishing, Social engineering, Human behaviour

1. Introduction

Just as junk postal mail is increasingly populating one's household letterbox, so it is with junk electronic mail (i.e. email) that arrives in one's personal or company email inbox. In both cases, much of this mail is advertising material and is seldom more than nuisance value. Examples of such emails include weight loss schemes, lonely heart offers and casino gambling. However, unlike junk postal mail, junk email has a much higher percentage of items that can be classified as nasty, sinister or offensive. These emails are almost always unwanted and anonymous and are known as 'rogue' emails. Examples of rogue emails include those embedded with a virus, those that contain pornographic or violent material and those that contain some form of malicious code like keystroke logging malware.

This paper is concerned with an additional type of rogue email that uses both social engineering and technical subterfuge to steal consumers' personal identity data and financial account credentials (Anti-phishing_Working_Group 2010) such as usernames, passwords, bank account and credit card details to enable the sender to obtain money fraudulently. These types of emails are the subject of this paper and they are known as phishing emails.

1.1. Justification for this Study

The Anti-Phishing Working Group reported 126,697 phishing email attacks in the second half of 2009, which is more than double the number of attacks recorded in the first half of 2009 (Aaron & Rasmussen 2010). It is difficult to determine the number and type of computer users that are susceptible to phishing attacks, or indeed, how much money has been lost by victims of such attacks. Notwithstanding, this ubiquitous problem is certainly serious enough to warrant this collaborative research project.

The impetus for this particular study originated from information system security professionals within the Australian Department of Defence who are well aware that to maintain a high level of security of their organisational systems requires more than just hardware, software and environmental controls. They appreciate that human aspects also play an important role in achieving this objective. More specifically, they believe that poor human behaviour, whilst using a computer, has the potential to negatively impact the security of organisational systems and data. Therefore they support the hypothesis that if computer user behaviour is improved then the level of security of organisational systems and data will increase.

Consequently, the Defence clients were interested in identifying factors that might predict injudicious behaviour. In particular, they were interested in demographic and personality characteristics of computer users that might be correlated with bad behaviour in regard to the management of emails, especially phishing emails. They felt that a preliminary study into a small but common area of computer usage such as responding to phishing emails may serve to highlight future research directions relating to human aspects of computer security and assurance.

1.2. Aim of this Paper

The aim of this paper is twofold. The first aim is to describe a scenario-based, role-play experiment that investigates how computer users respond to phishing emails compared to genuine emails. The second aim of this paper is to present the results of a data analysis and to discuss the impact that a selection of factors has on the behavioural responses of email users.

The structure of this paper is as follows. The next section summarises previous research that has been conducted and how this current research positions itself to address some of the apparent shortcomings and gaps in the literature. This is followed by an explanation of the research method and details about the survey questionnaire. Finally, the results are presented and discussed, limitations are conceded and conclusions are expressed.

2. Previous Research

Previous research on this topic can be broadly categorised into three types of study as follows.

2.1. Real Phishing Experiments

In these studies participants receive actual emails in their inbox and researchers observe how they interact with the various types of email. Usually, participants are not aware that they are participating in an experiment.

A real phishing study was conducted on a student population from the University of Indiana by Jagatic, Johnson, Jakobsson & Menczer (2007). They found that social context had a significant influence on the way in which participants responded. The authors also found that age and gender contribute to an individual's vulnerability, younger individuals were more susceptible, and females were more likely to click on the phishing link compared to males.

Very few studies of this type have been conducted because the ethical issues surrounding the deception and lack of consent makes it difficult to collect demographic data or conduct any cognitive or personality tests. For these reasons, the study described in this paper does not involve the sending of any emails. However, it attempts to emulate this process by presenting images of a large selection of emails and asking participants how they would respond if they received them in their inbox.

2.2. Phishing IQ Tests

In these studies, participants who are informed that they are participating in a phishing study, are presented with images of phishing and genuine emails, and are asked to make judgements concerning the authenticity of the images.

A study conducted by Furnell (2007) revealed that cues such as visual presentation, language use, and content influenced participant's decision making strategies. Generally, participants in the study were more likely to trust emails with the presence of visual factors such as banners and logos, and were less likely to trust emails with language errors or a forceful or urgent message. Results also indicated that participants were more likely to incorrectly classify genuine emails than phishing emails.

A phishing IQ test was also conducted by Dhamija, Tygar & Hearst (2006) who found visual presentation to be a powerful phishing tool. They also found that participants often ignored standard security indicators such as the address bar and status bar. Although this research provided valuable findings, the study was limited by its small participant pool of 22. In addition, only 20 website images were presented for participants to assess. In comparison, the study described in this paper presents 50 images for assessment by 117 participants.

Phishing IQ tests been criticised for lacking real-world validity. Essentially, in experiments where individuals are informed that they are undertaking a phishing study, this knowledge may make them more suspicious and in turn result in a bias towards 'phishing' decisions (Anandpara, Dingman, Jakobsson, Liu & Roinestad

2007). Since this level of suspicion may not be present when individuals are checking their personal inboxes (Furnell 2007), some researchers have questioned whether phishing IQ tests are able to measure susceptibility to phishing emails. This cognitive bias is generally referred to as the subject expectancy effect and needs to be considered when using a phishing IQ test design (Anandpara *et al* 2007). This phenomenon has often been overlooked by phishing researchers. Consequently, the study described herein, deliberately set out to have a control group (i.e. participants that were not informed that this was a phishing study) and a treatment group (participants that were informed).

2.3. Scenario-based Experiments

In these types of study (also known as role-play experiments) participants are presented with images of emails and then asked, often in person, how they would respond if they received such an email. Participants may or may not be informed that they are participating in a phishing study.

A scenario-based experiment was conducted by Downs, Holbrook & Cranor (2007) who found that participants who had a greater understanding and more experience with the internet environment were less susceptible to phishing attacks. They also found that perceived consequences were not a predictor of behaviour responses. Although this study had an impressive sample size of 232, participants were only exposed to five email images which limited the scope of their findings. The study described herein presents 50 email images covering activities such as shopping, banking and social networking.

Another study conducted by Downs, Holbrook and Cranor (2006) also used the scenario-based approach and included qualitative interviews about computer security and trust. These authors found that previous exposure to particular phishing scams reduced susceptibility and participants were particularly vulnerable to unfamiliar scams. Although this research provided some interesting insights, the authors (Downs *et al* 2006) acknowledge that the findings were significantly limited by a small sample of 20 participants.

Finally, Sheng, Holbrook, Kumaraguru, Cranor & Downs (2010) applied a scenario-based design to explore the relationship between demographic factors and phishing susceptibility. An online survey was completed by 1001 participants and it was found that two demographic factors; gender and age, were related to an increase in susceptibility to phishing attacks. Females were most vulnerable, in conjunction with participants aged between 18-25 years.

The study described in this paper was designed to address the shortcomings of this previous research into phishing emails by having a significant number of participants assess a comparatively large number of emails. Also, the study addressed the subject expectancy effect by informing half of the participants that they were participating in a phishing study.

3. Method

3.1. Overview

The aim of this research was to investigate the behaviour response of computer users when faced with phishing emails and to compare this to their response when having to deal with genuine emails.

This current study was a scenario-based role-play experiment that involved the development of a web-based questionnaire that was only accessible from a specific University computer laboratory. To prevent prior exposure of participants, it could not be accessed remotely and the web address was changed after every session.

Each session took less than one hour and was facilitated by a member of the research team who explained the study and remained on hand during the study to answer queries.

The participants were given different instructions depending on which session they attended. In the sessions held in the first week, participants were not informed that that they were participating in a phishing study but were told only that it was a study about how they manage emails. In the second week of sessions, participants were informed that this was a phishing study and a brief explanation of 'phishing emails' was given. This approach enabled the research team to measure the subject expectancy effect on participants and it also ensured that participants in the first week were not alerted by previous participants that this was a phishing study.

3.2. Subjects

The survey participants were students enrolled at the University of Adelaide. Two cohorts of students were invited to participate. The first cohort was recruited from the business school and the second cohort was recruited from the psychology school. A total of 117 students, consisting of 64 business students and 53 psychology students, attended one of the 16 scheduled sessions and completed the survey.

3.3. The Questionnaire

The questionnaire consisted of 3 sections. The first section contained the assessment of 50 email images (pdf files), all collected over time by the research team or found on the Internet. Half of these images were genuine emails and half were phishing emails. They were presented in random order. The 50 emails were carefully selected to ensure that they represented a range of topics that one would expect to see in an email inbox, such as shopping, banking and social networking. Each email image was modified to show that it was sent to 'Sally Jones', a fictitious character. Each participant was instructed to play the role of an advisor to Sally Jones by recommending how she should manage each of the emails.

The second section contained demographic questions about age, gender, education level and area of study, previous and current employment, spoken language and familiarity with computers.

The third section contained a personality test and a cognitive reflection test. The personality test was the public domain instrument known as the Big Five Inventory Scales (BFI) test (John & Srivastava 1999).

These three sections of the questionnaire were sequenced in this way to ensure that the demographic questions did not alert participants that they were involved in a phishing study until after they had completed their assessment of the 50 emails.

3.4. Variables Analysed

The following variables were subjected to analysis.

3.4.1. Familiarity with Computers

Participants were not asked, in a direct sense, to rate their familiarity with computers and so a variable was 'created' using the Predictive Analytics SoftWare (PASW) syntax coding to average the score for the following 24 demographic questions for each participant:

- How frequently do you use a computer from the following locations? (Home, Work, University, Other public computers)
- How frequently do you access the internet from the following locations? (Home, Work, University, Other public computers)
- How frequently do you access your email from the following locations? (Home, Work, University, Other public computers, Other private computers)
- How frequently do you engage in the following computer activities? (Email, Web surfing, Research, Word processing, Games)
- How frequently do you use the following applications? (PayPal, eBay, FaceBook, MySpace, Twitter, Online purchasing)

Each of these 24 questions required a forced likert scale response ranging between Daily (score = 1), 2-3 times a week (score = 2), Once a week (score = 3), Once a month (score = 4), Less than once a month (score = 5) and Never (score = 6). For each participant, the 24 likert scales were reversed, then aggregated and the total divided by 24 to give an average score between 1 and 6 whereby the higher the score, the more familiar they are with computers.

3.4.2. Personality Traits

Participants also completed the Big Five Inventory Scales (BFI) test (John *et al* 1999). This public domain measure consists of 44 items and provides data on the big five personality traits, namely, extraversion, agreeableness, conscientiousness, neuroticism and openness. This measure was included in the study to indicate if

specific personality characteristics or traits had any bearing on an individual's response to phishing emails. This test has been shown to have good psychometric properties when compared to other more comprehensive personality tests (John, Donahue & Kentle 1991; John, Naumann & Soto 2008; John *et al* 1999).

3.4.3. Cognitive Impulsivity

Participants completed the Cognitive Reflection Test (CRT), which is a very quick and efficient measure of cognitive impulsivity (Frederick 2005). The test consists of three questions and, in order to obtain a correct response, participants need to be able to stop and consider the question before providing an answer. The most obvious response is not the correct response, and a higher score relates to a better ability to control impulsivity. This particular test was selected because findings indicate that the predictive validity of this measure was equal or above other cognitive tests that claim to measure cognitive impulsivity (Frederick 2005). This measure was included in this study to investigate whether individuals who are more impulsive in their decision making are more likely to misclassify a phishing email as genuine.

3.4.4. Behavioural Response

Behavioural response refers to how participants responded to the following question for each of the 50 emails:

"How would you manage this email?

a) *Leave the email in the inbox and flag for follow up*
b) *Leave the email in the inbox*
c) *Delete the email*
d) *Delete the email and block the sender"*

For each participant, two behavioural response variables were calculated in the PASW dataset - a phishing email behaviour score and a genuine email behaviour score, as follows:

If the email being evaluated was one of the 25 genuine emails, the participant would score 2 for response a), 1½ for response b), 1 for response c) and ½ for response d). If the email being evaluated was one of the 25 phishing emails, the participant would score ½ for response a), 1 for response b), 1½ for response c) and 2 for response d). This scoring method enables a maximum score of 50 and a minimum score of 12.5 for each type of email.

The results of this analysis are described below.

4. Results

Table 1 below shows the mean scores for managing phishing, genuine and all emails and indicates that participants who were informed that this was a phishing study

managed phishing emails much better than those that were not informed. There was very little difference between the two groups of participants in managing genuine emails, although overall, genuine emails were managed better than phishing emails.

		Not informed	Informed	Total
Mean behavioural score	**Phishing emails**	27.3 *(SD=6.02)*	33.3 *(SD=6.20)*	30.3 *(SD=6.79)*
	Genuine emails	35.0 *(SD=4.55)*	36.2 *(SD=3.85)*	35.6 *(SD=4.24)*

Table 1: Mean behavioural scores for managing emails

In order to investigate the factors that may affect an email user's response to phishing emails compared to genuine emails, a correlation analysis was performed to explore the interrelationships among a number of variables in the PASW dataset.

The tables below show the relationship between each of seven variables with the behavioural response in managing emails by participants who were not informed that this was a phishing survey (Table 2) and those that were (Table 3). Spearman's correlation coefficient (rho) was used to indicate the strength of the relationships and the 2-tailed p value indicates how much confidence people should have in the results. The variable N indicates the number of participants in each group.

	Behavioural Response with Genuine Emails		Behavioural Response with Phishing Emails	
	rho	**p (2-tailed)**	**rho**	**p (2-tailed)**
Familiarity with Computers	0.039	0.767	0.066	0.618
Extraverted	0.014	0.914	0.206	0.118
Agreeable	-0.023	0.861	-0.137	0.300
Conscientious	-0.075	0.570	0.005	0.970
Neurotic	-0.069	0.604	0.055	0.679
Open	-0.010	0.939	0.249	0.057
Cognitive Impulsivity	-0.063	0.635	0.215	0.101

Table 2: Not informed that this was a phishing survey (N = 59)

	Behavioural Response with Genuine Emails		Behavioural Response with Phishing Emails	
	rho	p (2-tailed)	rho	p (2-tailed)
Familiarity with Computers	-0.031	0.819	0.315	0.016
Extraverted	0.104	0.436	-0.107	0.425
Agreeable	-0.240	0.069	0.102	0.447
Conscientious	0.041	0.762	0.175	0.189
Neurotic	0.005	0.973	-0.062	0.646
Open	0.089	0.506	0.113	0.400
Cognitive Impulsivity	0.149	0.265	0.032	0.812

Table 3: Informed that this was a phishing survey (N = 58)

5. Discussion

5.1. Familiarity with Computers

For the informed participants, that is, those who were told before the survey that "we were specifically interested in your ability to identify phishing emails", it was found that the more familiar people were with computers, the better they managed phishing emails. This finding was statistically significant and somewhat predictable because computer-savvy people are likely to be more aware of the risks and consequences associated with phishing emails than people who don't use computers very often.

For the not-informed participants, that is, those who were not told that this was a phishing study, their familiarity with computers had no significant effect on how they managed phishing emails.

In terms of managing genuine emails, the variable, familiarity with computers, had no significant impact.

5.2. Personality Traits

The analysis of scores from the Big Five Inventory Scales (BFI) personality test indicated that for the not-informed participants, the more extraverted they were, the better they managed phishing emails. This was a significant association. This same observation applied to the 'openness' personality trait such that the more open they were, the better they managed phishing emails. Both of these findings are interesting because they appear to be counter-intuitive since one would assume that extraverted and open people are likely to be more trusting of others which would mean treating some phishing emails as genuine.

In terms of managing genuine emails, personality trait variables had no significant impact except for 'agreeableness', which had a negative effect on informed participants. The more agreeable they were, the worse they managed genuine emails. However, this same observation did not translate to phishing emails.

5.3. Cognitive Impulsivity

For not-informed participants, it was found that the less impulsive people were (i.e. those with a higher CRT score), the better they managed phishing emails, although this was a marginally significant result. In other words, those that tended to deliberate over a phishing email appeared to respond better. This finding agrees with Kumaraguru, Rhee, Sheng, Hasan, Acquisti, Cranor & Hong (2007) who found that participants with higher CRT scores were less likely to click on the phishing emails, although the results in their study were not statistically significant either.

The cognitive impulsivity of informed participants had no significant impact on how they managed both phishing and genuine emails.

6. Limitations

This study does not claim to evaluate actual susceptibility to phishing emails. There are several reasons for this. Firstly, this was a role-play experiment that presented images of emails to participants and prompted them to indicate what they would do with them if they actually arrived in their inbox. Consequently, participants were not required to, nor could they, click on embedded links, delete the email, flag it for follow-up or block the sender. A second reason why this study is not attempting to evaluate susceptibility to phishing emails is that there were no consequences as a result of the action that participants indicated they would take. In other words, participants would have been aware that there was nothing to lose by suggesting the wrong action to take, or conversely, there was nothing to gain by suggesting the correct action. Additionally, participants in this study did not have the contextual information that would often influence decision-making. For example, they did not know whether 'Sally Jones' was a member of a certain bank, social networking site or purchasing website.

The 117 participants in this study were all University of Adelaide students - 64 enrolled in a business course and 53 enrolled in a non-business course. A sample such as this is likely to be biased in terms of age, education, familiarity with computers and various other factors. Therefore, the conclusions that follow are not necessarily representative of the general population of email users.

7. Conclusion

The aim of this research was to investigate the behaviour response of computer users when phishing emails arrive in their inbox and to compare this to their response when they receive genuine emails.

The findings indicate that genuine emails were managed better than phishing emails whether or not participants were informed beforehand that this was a phishing study.

Although informed participants performed better than not-informed participants in managing both types of email, they were significantly better in managing phishing emails. This implies that computer users should be continually reminded, via security awareness sessions and other risk communication practices, that phishing emails are a serious threat to information security.

This research also analysed the impact that a selection of factors has on email users when they are faced with having to deal with phishing emails. The survey results indicate that cognitive impulsivity and the personality traits of extraversion and openness are each associated with improved performance in managing phishing emails, by the participants who were not informed that this was a phishing study. Interestingly, these factors did not have a significant impact when participants were informed but their familiarity with computers did have a positive effect in managing phishing emails. This implies that computer-savvy participants needed reminding that phishing emails must be managed carefully. It was also observed that the personality trait of agreeableness had a negative effect on how informed participants managed genuine emails. This could be explained as an 'if in doubt, treat as a phishing email' response.

In the opinion of the authors, these results, although inconclusive, satisfy the aims of this research, which was to indicate whether further investigation was warranted into the impacts of personal characteristics on computer user behaviour in other areas of information technology usage. For example, it would be interesting to examine how various types of people manage passwords or use the Internet. The outcomes of such research would assist management in their endeavours to improve computer user behaviour and, as a result, help to mitigate risks to their organisational information systems, thus making them more secure.

8. References

Aaron, G & Rasmussen, R 2010, Global Phishing Survey: Trends and Domain Name Use in 2H2009, AntiPhishing Working Group (APWG), Lexington, MA, USA.

Anandpara, V, Dingman, A, Jakobsson, M, Liu, D & Roinestad, H 2007, 'Phishing IQ Tests Measure Fear, not Ability', Financial Cryptography and Data Security, pp. 362-366.

Anti-phishing_Working_Group 2010, Phishing Activity Trends Report 2nd quarter 2010, APWG, pp. 1-11, viewed 3rd March 2011 <http://www.antiphishing.org/reports/apwg_report_q2_2010.pdf>.

Dhamija, R, Tygar, J & Hearst, M 2006, 'Why Phishing works', SIGCHI Conference on Human Factors in Computing Systems, Montreal, Quebec, Canada, pp. 581-590.

Downs, J, Holbrook, M & Cranor, L 2006, 'Decision Strategies and Susceptibility to Phishing', Second Symposium on Usable Privacy and Security (SOUPS), Pittsburgh, Pennsylvania, vol. 149, pp. 79-90.

Downs, J, Holbrook, M & Cranor, L 2007, 'Behavioral Response to Phishing Risk', Proceedings of the anti-phishing working group's 2nd annual eCrime researchers summit (eCrime '07), New York, NY, USA, pp. 37-44.

Frederick, S 2005, 'Cognitive Reflection and Decision Making', Journal of Economic Perspectives, vol. 19, no. 4, pp. 25-42.

Furnell, S 2007, 'Phishing: can we spot the signs?', Computer Fraud & Security, vol. 2007, no. 3, pp. 10-15.

Jagatic, T, Johnson, N, Jakobsson, M & Menczer, F 2007, 'Social Phishing', Communications of the ACM, vol. 50, no. 10, pp. 94-100.

John, O, Donahue, E & Kentle, R 1991, The Big Five Inventory-Versions 4a and 54, University of California, Institute of Personality and Social Research, Berkeley.

John, O, Naumann, L & Soto, C 2008, 'Paradigm Shift to the Integrative Big Five Trait Taxonomy: History, Measurement and Conceptual Issues' in Handbook of Personality: Theory and Research (3rd ed.), O John, R Robins & L Pervin (eds), Guilford Press, New York, USA, pp. 114-158.

John, O & Srivastava, S 1999, 'The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives' in Handbook of personality: Theory and research (2nd ed.), L Pervin & O John (eds), Guilford Press, New York, USA, pp. 102-139.

Kumaraguru, P, Rhee, Y, Sheng, S, Hasan, S, Acquisti, A, Cranor, LF & Hong, J 2007, 'Getting users to pay attention to anti-phishing education: Evaluation of retention and transfer', paper presented to 2nd Annual eCrime Researchers Summit, Pittsburgh, PA, USA, October 4-5.

Sheng, S, Holbrook, M, Kumaraguru, P, Cranor, LF & Downs, J 2010, 'Who Falls for Phish?: A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions', paper presented to The 28th ACM Conference on Human Factors in Computing Systems (CHI 2010), Atlanta, Georgia, USA, April 10-15.