International Journal of Information & Network Security (IJINS)
Vol.1, No.4, October 2012, pp. 235~242
ISSN: 2089-3299

Evaluation of Field Phishing Study Setup Method

Yunsang Oh*, Takashi Obi**
* Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology
** Imaging Science and Engineering Laboratory, Tokyo Institute of Technology

Article Info

Article history:

Received July 10th, 2012
Accepted July 28th, 2012

Keyword:

Phishing
Experiment
Experimental Ethic

ABSTRACT

Phishing threatens to topple information societys stability because this erodes trust in its underlying infrastructure. From the insight, researchers are attempting to quantify how people fall for deceit. However, in-lab studies are challenged with ecological and external validity issues. So researchers conducting security usability studies are engaged in deceit-based field studies of users that are conducted without prior consent. Unfortunately, such studies can expose researchers to ethical risks since field studies usually mimic real phishing. Here, we present studies about how researchers managed risks for previous deceit-based studies, not only in usable security but also in other research areas such as psychology, and then propose and evaluate recommendable experiment design method and ethic guideline for ethical and valid phishing study.

Corresponding Author:

Yunsang Oh,
Interdisciplinary Graduate School of Science and Engineering,
Tokyo Institute of Technology,
4259-R2-60 Nagatsuta, Midori-ku, Yokohama 226-8503, Japan
Email: oh.y.ab@m.titech.ac.jp

1. INTRODUCTION

To fight against phishing effectively, a number of studies have been done previously to identify threat models about users' behavioral response to phishing. However, these and other recent phishing-related laboratory studies are not readily generalizable to a larger population and their authors were able to find few, if any, correlation between demographics, personal characteristics, and behaviors relevant to the studies. While researchers in other areas of computer science are able to rely on controlled, in-lab studies of solicited users, this approach is not as favored by those studying phishing.

In this sense, phishing researchers recently favor deceit-based field studies of users that are conducted without their knowledge or prior consent. In general, these field studies can be conducted by mimicking real phishing attacks for the experiments' effectiveness. However, this makes burdens to experiment since individual participants can feel the experiment is unethical and they did not consent to the experiment. Even worse, researchers can be exposed to legal risks.

In order to overcome the ethical and legal setbacks for valid phishing studies, we need to investigate previous works about how they managed the issues. We found some meaningful discussions about how to design phishing experiments while reducing ethical and legal risks; researchers already conducted field experiments and clarified how they managed the risks. In this paper, we contribute in the following areas. First, we surveyed what previous field studies mention about designing a deceit-based field experiment (see section 3). Then, we propose an experiment setup tutorial considering surveyed issues. In addition, we will present an ethic guideline for researchers planning deceit-based phishing field experiments. For this, we will focus on legal situations in Asian countries such as in Japan, South Korea, and the United States (see section 4). Finally, we will evaluate our proposal (see section 5). Despite this paper presents a tutorial to help researchers concerning ethical or legal troubles, this information must not be taken as legal advice.

2. PHISHING EXPERIMENTS METHODS OVERVIEW

Users' perception process toward phishing sites may vary depending on services' characteristics. Financial and payment service authenticity can be effectively faked because users are trained to authenticate the services without relevant vulnerability investigation due to the services' trustworthiness. Therefore, simple and uniform Web page design can be effectively exploited to lead users to ignore phishing detection signals in Web browsers [1, 11]. Thus far, a variety of tools and techniques are advancing the fight against phishing. However, phishing detection remains an arms race. All of defending tools and techniques stand to benefit from knowledge of users' behavioral response to phishing. For that, researchers have used a variety of methods: Survey, Lab Studies, and Field Studies in user studies designed to gain the user behavior model into the issues [5].

2.1. Survey

"Survey" is a study where researchers collect data by asking people questions. This can be conducted using paper questionnaires, email, web-based survey forms, or occasionally by telephone or in public spaces. Survey has widely been conducted to understand users' mental models and decision processes when they faced phishing. However, surveys tend to underestimate damages; many victims are unaware that an attack occurred or are unwilling to disclose that they fell for it.

2.2. Lab Study

"Lab Study" is generally conducted by a role playing in order to test users' susceptibility to phishing attacks and evaluate the effectiveness of anti-phishing toolbars and training materials. Generally, participants play a fictitious role and use personal information associated with that role. Lab studies are very helpful in understanding user behavior in a given situation. However, this study method has tradeoffs and faces validity challenges: most of these studies are challenged with ecological (whether the methods, materials, and settings are similar to real life) and external (whether the results are generalizable) validity issues. It overestimate attack awareness because of expectancy bias - the mere knowledge of the study's existence biases its likely outcome.

2.3. Field Study

"Field Study" can be used to evaluate participants' susceptibility to phishing with higher confidence than "lab studies", but not to evaluate the effectiveness of training or anti-phishing tools. Experiment by this method mimics real phishing and is conducted without participants' consents. Despite of ethical and legal issues, this method is most effective to understand precise users' behavioral responses against phishing attempts and to find an unidentified new threat model. To minimize a disparity between participants' actions in the lab study and "would be" actions in everyday computer practices, the field study seems inevitable. Especially, for security usability research in general, validating findings of the previous lab study result is recently populated with the contribution of finding evidence of a strong bias in the lab environment [10].

3. LESSONS FROM PREVIOUS STUDIES

In section 2, we introduced three phishing experimental methods. In practice, setting up survey and lab study can be conducted without ethical and legal issues. However, in order to understand various aspects of users' susceptibility to phishing attacks precisely, deceit-based field study is highly required in order to defeat phishing despite of risks. In this section, we will introduce what we learned from the previous studies. Referring to previous works will contribute to propose a tutorial for designing phishing field experiments.

3.1. Ethics for Deception Based Research in General

Some methodologies in social research, especially in psychology involve deception, where researchers purposely mislead or misinform the participants about the true nature of the experiment. The rationale for deception is that humans are sensitive to how they appear to others (and to themselves) and this self-consciousness might interfere with or distort from how they actually behave outside of a research context (where they would not feel they were being scrutinized). In the past, some psychological experiments question about the deception based research ethics, because study participants were not aware of accurate information about the study and they did not inform their consents. Critics argued that participants could suffer emotional stress.

In an experiment conducted by Milgram [8], the researchers told participants that they would be participating in a scientific study of memory and learning. In reality, the study measured the willingness of study participants to obey an authority figure who instructed them to perform acts that conflicted with their

personal conscience. Specifically, the researchers monitored how ordinary people, simply doing their jobs, and without any particular hostility on their part, can become agents in a terrible destructive process, that was observed from Nazi during the world war II. The Milgram's experiment raised questions about the research ethics of scientific experimentation because of the extreme emotional stress suffered by the participants.

Through their participation in the experiment, many participants realized that they were capable of committing acts of extreme violence on other human beings. After having this realization, many subjects experienced prolonged symptoms of anxiety [7]. Despite of Milgram's academic contributions with this experiment, a psychological study like this would not be allowed in most countries today, due to ethical considerations.

Currently, psychologists in many countries follow the experiment ethics guidelines, such as one created by American Psychological Association (Ethical Principles of Psychologists and Code of Conduct, http://www.apa.org/ethics/code/index.aspx). These guidelines mention how to handle deceit based experiments. However, though commonly used and allowed by the ethical guidelines, there has been debate about whether or not the use of deception should be permitted in psychological research experiments.

3.2. Legal Risks in Phishing Study

Legal risks in deceit-based experiments are the most critical issue for researchers, and many of them face difficulties in legal consulting. While it is unlikely that suits would be brought, the legal risks definitely exist. Firstly, researchers can be accused of sending phishing e-mails by subjects. In South Korea, the "Act on Promotion of Information and Communication Network Utilization and information Protection" prohibits a behavior to harvest personal data by deceiving or to lead victims to provide personal data. Japanese penal code also clearly prohibits a fraud, and strictly restricts spam messages which can be exploited for phishing. Also taking sensitive personal data such as password or medical information is strictly controlled. Copyright law can also be applied to prohibit the public transmission of the copyrighted website duplication by the third party. Secondly, even if no one is harmed, or can claim they were adversely affected, the law still permits the third-party to accuse researchers of phishing experiments. This has variants depending on locations and countries. As we mentioned in Table 1, even in the United States, Californian law permits the attorney general to bring a civil action against the sender of phishing email unlike to the other states. Therefore it is not recommended for researchers located in California to conduct real phishing experiments [9].

To the best of our knowledge, Soghoian's work [9] was the first contribution to propose the best practices for designing a field phishing experiment with the considerations of legal issues. In Table 1, we summarize Soghoian's best practices that may help researchers avoid running afoul of the law.

Table 1. Guideline for Phishing Experiment Design (Proposed by Soghoian [9])

Avoid California
Researchers who are located within the state of California should probably not engage in phishing field experiment. Researchers should also avoid the experiment targeting people located in the state. The law permits the Attorney General to bring a civil action against phishing email senders.
Terms of Service and User Accounts
Researchers should avoid the automated scraping of sites that require users first create and login to an account. Usually, terms of service are required to be agreed by users for the account creation.
Application Programming Interfaces
If researchers require the automated collection of data, use the site's public Application Programming Interface (API).
Moderation
When scraping a website, researchers should make sure to limit the number of requests in order to maintain a upper limit to the bandwidth used.
Discretion
Companies' reputation must be protected. Their name should be hided by doing obfuscation or vagueness.
Use Caches
It is better to scrap website contents cached by third party than scrap directly.
Commerce
Researchers must avoid any possible personal profit from the phishing experiment. Remove advertising banners from the research website or blog.

3.3. Participants' Responses Analysis in Previous Phishing Studies

3.3.1. Criticism

Jagatic et al. [4] conducted a field study to identify a new phishing threat model that attackers exploit publicly available personal information on social networks. With a real phishing experiment, this work succeeded to identify the threat model named "Social Phishing". Specifically, this work was the first to harvest voices from experiment participants when they understood they participated in the experiment without their consent. Participants' opinions were collected through the research blog.

Despite of meaningful academic contribution, the researchers failed to clarify all ethical issues. They conducted a real phishing by sending an email to some students in a university requesting to visit a URL embedded in the message. The email was delivered from the researchers to the recipients with the name of the sender whose name was taken from Internet social network services. About the experiment, there were not much complaints, but some participants complained that it was unethical, inappropriate, illegal, unprofessional, fraudulent, self-serving, and/or useless. They also called for the researchers to be fired, prosecuted, expelled, or reprimanded.

3.3.2. Attack Timing

In Jagatic's work [4], the attack timing may be one of the reasons about the harsh reactions of some participants. The researchers guessed that the participants' reaction is correlated with the attack timing. The experiment was carried out near the end of the semester. This may intensified the stress felt by some students. Ferguson [2] also pointed out the importance of attack timing. He conducted a real phishing experiment targeting students in a university. The phishing message included an embedded link to the phishing site and contents about troubles in student's grade reports. Actually this experiment was conducted three weeks before the end of the semester. Emails regarding exams and grades can get students' attention as the semester draws to a close.

3.3.3. Phishing Education in Organizations

Ferguson's work [2] was conducted to evaluate the effectiveness of security education in the campus. In practice, organizations such as universities or companies can set up a real phishing experiment to investigate the effectiveness of security education for students or employees. In addition, challenge to maintain a robust network against phishing threats is an attractive goal to network administrators. Therefore, mimicking a real phishing attack can be an appropriate approach to measure the knowledge and reaction of organizations' members to phishing attacks. This result will be referred to for creating an internal security policy or planning further education. For example, an experiment by Kumaraguru et al. [6] was conducted in a Portuguese company, because the company was primarily interested in studying the vulnerability of their employees towards phishing emails.

4. PROPOSED TUTORIAL

In the previous section, we summarized worth referring previous field studies. Despite of observed setbacks, no more delay should be allowed to prepare a safe experimental environment to fight against phishing attackers who threaten the entire trust of information society. Researchers already pointed out that ethical and critical issues should be carefully handled while setting up the experiments [3, 5] but they did not provide a clear approach such as tutorial. In this section, we introduce the experimental flow and ethical/legal considerations in each stage as an experiment setup tutorial in Table 2.

Table 2. Field Phishing Study Setup Check List

Steps	To Do List	Ethical and Legal Considerations
1	Phisher's Strategy Decision (Sect 4.1)	For the phishing message contents, avoid factors that are reasonably expected to cause severe emotional distress to participants.
2	Phishing System Implementation (Sect 4.2)	Do not harvest sensitive private data by the system.
3	Phishing Site and Email Building (Sect 4.3)	Do not violate local copyright law. Refer to the legal advices in Table 1.
4	Email Sender and Receiver Decision (Sect 4.4)	Do not intrude 3rd party systems illegally to harvest participant's personal data.
5	Phishing Experiment Execution (Sect 4.5)	Do not violate local anti-phishing or anti-spam laws.
6	Follow-Up (Sect 4.6)	

4.1. Phisher's Strategy Decision

Decide malicious strategies based on a hypothesis what characteristics of a target service brand impact users' decision for perceiving their trust. With the strategies, design a phishing website and a bogus e-mail including fake contents. However, the phishing website and e-mail must not contain ethically or legally improper contents to cheat experiment's participants, despite the more ethical and legal constraints may lead to the less reality in obtained experimental data. In this sense, we provide a minimum restriction what to avoid as followings. Our aim is to minimize factors that we reasonably anticipate to cause severe emotional distress.

Unfair Discrimination Researchers must not manipulate participants with unfair discrimination based on age, gender, gender identity, race, ethnicity, culture, national origin, religion, sexual orientation, disability, socioeconomic status, or any basis proscribed by law.

Harassment Researchers must not manipulate participants with harassment based on factors such as participants' age, gender, gender identity, race, ethnicity, culture, national origin, religion, sexual orientation, disability, language, or socioeconomic status.

Adult Contents For a phishing experiment, researchers must not manipulate participants with adult contents such as violence or pornography which are not allowed for minors.

Financial Data Request For a phishing experiment, researchers must avoid to lead participants to provide private financial data such as bank account or credit card numbers. Participants can be anxious about potential financial accident exploiting the submitted data. If accidentally collected such data, researchers must show evidences to participants that the data is neither recorded in the researcher's system nor transferred over the network in the plaintext.

Medical Treatment Information For a phishing experiment, researchers must not manipulate partici- pants with fake information about medical treatment. Phishing with medical treatment information can be effective to manipulate participants who are suffered with specific diseases, but such fake information can disappoint them seriously.

4.2. Phishing System Implementation

Set up an infrastructure for an attack mimicking a real phishing experiment: a Web server for a phishing website and a database to record user behaviors. The infrastructure should be cautiously designed not to remain sensitive private data submitted by experiment participants. In addition, secure the experimental system and database. Some participants, who noticed phishing experiment, can retaliate against the experimental system by modifying or removing harvested study data. This is a significant loss for researchers.

4.3. Phishing Site and Email Building

Design a phishing website and a bogus e-mail including fake contents. This can be achieved by two methods. Firstly, a fake website (not-existing website) can be created. This method is used when researchers need to demonstrate unidentified likely phishing threats. Secondly, an existing website can be duplicated. This method can be used when researchers need to identify vulnerabilities of an existing website. In this case, researchers require the company or organization's cooperation who is operating the website. Building the phishing website must not violate the copyright law. In addition, refer to other legal advices in Table 1.

4.4. Email Sender and Receiver Decision

A sender's name should be carefully decided to avoid troubles to people or organizations whose names and email had been spoofed as senders. Obfuscation or vagueness should be implemented if required. For receivers (experiment participants), a phishing message can target general users or a specific Internet user. In the former case, a phishing message can be reused for all participants. On the other hand, in the latter case, a phishing message may include participants' private data to bind the message to their real identity. For that, researchers need to harvest private information about the target users. However, illegal actions like intrusion to 3rd party services must not be conducted for the purpose.

In addition, randomness in selecting participants is important. Establishing a high degree of randomness involved selecting participants while avoiding the "high-beam" effect; the effect refers to the situation where drivers on the side of the highway spot a law enforcement officer at the side of the road looking to catch speeders. These drivers flash their high beam headlights to warn oncoming drivers that a law

enforcement officer is targeting speeders, encouraging them to slow down. Participants with a priori knowledge of the bogus e-mail would either deliberately handle the embedded link or totally ignore the message. Either way, the data would be skewed.

4.5. Phishing Experiment Execution

First, decide experiment timing before the execution. Depending on the timing when participants receive a phishing message, their responses would be biased. For example, students will react sensitively when they receive a phishing message about an exam just before the exam. Second, send the phishing message and monitor user behaviors. JavaScripts or PHP scripts can be used to monitor user behaviors such as clicking links, or to modify the phishing website dynamically for various scenarios. Before executing the experiment, fully understand local laws prohibiting spam e-mail or phishing itself. Restricting the phishing website visibility within participants only is a reasonable method to reduce legal responsibility.

4.6. Follow-Up

Researchers must receive consents from participants and optionally interview them after phishing experiment. Participant contact should be promptly conducted when analysis about the corresponding participant is completed. Interview may be required to explain about the research and to conduct surveys. Through the interview, user education against phishing threats can also be conducted to help them avoid potential phishing threats. If some participants are upset with the experiment, an interview is highly required to explain the research purpose and apologize. Lastly, researchers should permanently remove all remaining phishing websites when the experiment is over in order to avoid unexpected damages to stakeholders. This must be definitely conducted in case that the experimental phishing website is a duplication of an existing website.

5. EVALUATION OF PROPOSAL: A CASE STUDY

In 2010, we studied a threat model about how government's Web services like e-Government can be hijacked by phishing attackers [12]. Our goal was to identify some characteristics of the services and to investigate how phishers can effectively exploit the characteristics. In order to clarify the unexisting threat model, we needed to assess the new attack. So we adopted the approach to perform experiments that mimic real phishing attacks, thereby measuring the actual success rates by making sure that our study cannot be distinguished by the subjects from reality.

In this section, we evaluate how our proposal can be used in a real field phishing study. We will present what information we obtained from the experiment, and how we minimized experiment subjects' complaints.

5.1. Effectiveness of Using Webscript

By using Webscript language, our experimental system showed an alert to researchers immediately when subjects opened the experimental phishing message or interacted with the experimental phishing website. We linked webscripts to experimental phishing e-mail and website in order to monitor subjects' behaviors like button clicks, and we could effectively find a subject contact timing in a timely manner.

With this functionality, we could minimize the time duration between the timing when the experiment subjects become aware of the phishing and the timing of researcher's contact by monitoring subjects' interactions with the phishing site. If contact is delayed, subjects may complain about the experiment to the corresponding institutes who handles cyber crimes. They also forget their psychological state if the contact is delayed. In this case, we cannot observe a meaningful research result.

We could made a phishing website invisible from others except experiment subjects by using Webscript. This help avoid complaints and troubles from third parties. For that, we embedded an invisible random number in each phishing email and let the phishing website be opened only when the embedded random number was authenticated. We could also observe subjects' behaviors toward well-known security threats in Webservice environment such as malware download or XSS attack.

5.2. Importance of Interview

We found that some subjects were upset about our phishing experiment with the following reasons. Some were disappointed about themselves since they were not careful toward phishing threat. In this case, they appreciated us because the experiment reminded them phishing threat and carefulness. With the interview, we could apologize about the experiment and conduct an education for subjects about how to avoid phishing attack and why the privacy control function in the commercial social network services in Web is important.

Our immediate contact and interview to experimental subjects helped us observe the subjects' psychological state toward a real phishing in a timely manner; we successed to interview subjects before they forget what they feel about the phishing.

5.3. Ethic and Legal Issues

We had no ethical or legal trouble with experiment subjects'.

6. CONCLUSION

Phishing, the automated type of fraud is a relatively recent phenomenon. It becomes a social problem of quite catastrophic dimensions. This deceit-based attack threatens the web services' success because this erodes trust in its underlying infrastructure. It is in this spirit that we must identify possible security breaches caused by human factors and then propose defenders' strategies that mitigate phishing threats.

As a result of these insights, an increasing number of researchers and practitioners are attempting to quantify risks and degrees of vulnerabilities in order to understand where to focus protective measures. "Field study" is recently favored by many researchers, because we can obtain the most precise user behavior models. But it has ethical and legal setbacks to overcome because it mimics a real phishing. In order to reduce burdens to researchers, we need to establish a legal policy and advice to protect researchers. Unfortunately, to the best of our knowledge, no countries still have such consideration for phishing researchers.

In this paper, we propose a tutorial for the field phishing experiment setup built by referring previous phishing experiments. We also evaluated the effectiveness of the proposal by a case study; we presented our experiment based on the proposed tutorial to identify the phishing threat model targeting Web services.

These technical implementations cannot fully guarantee the legally safe experiment, but can be a good complement to minimize risks. Despite our contribution cannot be a legal advice, we aim to start related discussions to fight against phishing threats. In addition, our contribution must not be exploited as a legal advice for non-academic purposes.

REFERENCES

[1] Rachna Dhamija, J. D. Tygar, and Marti Hearst. "Why phishing works." In CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems, pages 581–590, New York, NY, USA, 2006. ACM.

[2] Aaron J. Ferguson. "Fostering e-mail security awareness: The west point carronade." EDUCAUSE QUARTERLY, pages 54–57, 2005.

[3] Peter Finn and Markus Jakobsson. "Designing and conducting phishing experiments." In IEEE Technology and Society Magazine, Special Issue on Usability and Security. IEEE, 2007.

[4] Tom N. Jagatic, Nathaniel A. Johnson, Markus Jakobsson, and Filippo Menczer. "Social phishing." Commun. ACM, 50(10):94–100, 2007.

[5] Markus Jakobsson, Nathaniel Johnson, and Peter Finn. "Why and how to perform fraud experiments." IEEE Security and Privacy, 6(2):66–68, 2008.

[6] Ponnurangam Kumaraguru, Steve Sheng, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong. "Lessons from a real world evaluation of anti-phishing training." In Proceedings of the The Third Anti-Phishing Working Group eCrime Researchers Summit, eCrime Researchers Summit '08, 2008.

[7] Robert J. Levine. "Ethics and regulation of clinical research." Yale University Press, 1988.

[8] Stanley Milgram. "Behavioral study of obedience." In the Journal of Abnormal and Social Psychology, 67(4):371–378, October 1963.

[9] Christopher Soghoian. "Legal risks for phishing researchers." In Proceedings of the The Third Anti-Phishing Working Group eCrime Researchers Summit, eCrime Researchers Summit '08, 2008.

[10] Andreas Sotirakopoulos, Kirstie Hawkey, and Konstantin Beznosov. "On the challenges in usable security lab studies: Lessons learned from replicating a study on ssl warnings." In Proceedings of the 8th Symposium on Usable Privacy and Security (SOUPS) 2011, 2011.

[11] Min Wu, Robert C. Miller, and Simson L. Garfinkel. "Do security toolbars actually prevent phishing attacks?" In Proceedings of the SIGCHI conference on Human Factors in computing systems, CHI '06, pages 601–610, New York, NY, USA, 2006. ACM.

[12] Oh Y, Obi T, Lee JS, Suzuki H, Ohyama N. "Empirical analysis of Internet identity misuse: case study of South Korean real name system." In Proceedings of the 6th ACM workshop on Digital identity management (DIM '10). New York, NY, USA, 2010. p. 27–34.

BIOGRAPHY OF AUTHORS

Yunsang Oh received the B.S. and M.S. degrees in Computer Science and Engineering from Sogang University, Korea in 1997 and 2002, respectively. During 2003-2009, he worked in Samsung Electronics. He also served a member of OMA for DRM standardization in mobile networks. He is currently studying at Tokyo Inst. of Tech, Japan as a Ph.D. student.

Takashi Obi received his M.E. and Ph.D. degree in information processing from Tokyo Inst. of Tech, Tokyo, Japan, in 1992 and 1997, respectively. He was a Research Associate (1995-2002) at Imaging Science and Engineering Lab., Tokyo Inst. of Tech. and a Visiting Researcher (1998, 2000) at Univ. of Pennsylvania. He has been an Associate Professor at Dept. of Information Processing, Tokyo Inst. of Tech. since 2002.