The African Journal of Information Systems

Volume 11 | Issue 3 Article 2

7-1-2019

A Naturalistic Methodology for Assessing Susceptibility to Social Engineering Through Phishing

Paula Musuva
United States International University - Africa, pmusuva@gmail.com

Christopher Chepken
University of Nairobi, chepken@uonbi.ac.ke

Katherine Getao
Ministry of Information, Communications and Technology, kate.getao@gmail.com

The African Journal of Information Systems

A Naturalistic Methodology for Assessing Susceptibility to Social Engineering Through Phishing

Research Paper

Volume 11, Issue 3, July 2019, ISSN 1936-0282

Paula M. W. Musuva
United States International University - Africa
pmusuva@gmail.com

Christopher K. Chepken
University of Nairobi, Kenya
chepken@uonbi.ac.ke

Katherine W. Getao
Ministry of ICT, Kenya
kate.getao@gmail.com

(*Received May 2018, Accepted December 2018*)

ABSTRACT

Phishing continues to be a prevalent social engineering attack. Attacks are relatively easy to setup and can target many people at low cost. This study presents a naturalistic field experiment that can be staged by organisations to determine their exposure. This exercise provides results with high ecological validity and can give organisations the information they need to craft countermeasures to social engineering risks. The study was conducted at a university campus in Kenya where 241 valid system users, also known as "insiders," are targeted in a staged phishing experiment. The results show that 31.12% of the insiders are susceptible to phishing and 88% of them disclose passwords that grant access to attackers. This study outlines various ethical considerations that ensure such exercises do not present any actual harm. The design of data collection instruments is discussed in depth to allow organisations the opportunity to develop similar tools for routine threat assessment.

Keywords

Social Engineering, Phishing, Unintentional Insider Threat, Threat Assessment, Naturalistic Methodology, Information Security.

INTRODUCTION

Social engineering is the use of manipulation by malicious outsiders to get unsuspecting insiders to compromise an organization's information security by providing access to confidential information or protected information systems (Luo, Brody, Seazzu, & Burd, 2011). One prevalent type of social engineering is phishing. Social engineering through phishing is a type of unintentional insider threat.

The term insider is used to refer to authorized users of information systems who operate within an organization's trust boundaries. These insiders often pose as information security threats when they accidentally expose their systems to attack. This is referred to as the unintentional insider threat (CERT, 2013).

Phishing is described by the Anti-Phishing Working Group (APWG, 2018) as a criminal attack that uses deception over a technical medium in order to get users to give out their personal data, login credentials and other confidential information. The deception aims at getting the user to think that the communication is a legitimate request for their confidential data or system access. Another way to describe phishing is simply 'fishing' for data (James, 2005). This is the use of social deception (the fishing bait) with the aid of communication technologies such as apps, email or websites (the fishing rod) to compromise the security of an information system (the catch).

Background

The most common technique for delivering phishing attacks is email because it provides a way to reach large numbers of people with little effort and low cost (APWG, 2018; James, 2005; Kumaraguru, Rhee, Acquisti, et al., 2007). In addition, once an email is delivered to an insider's inbox, it is considered to have crossed the external perimeter defenses and is now inside an organization's network. This makes it a very effective way of compromising information systems from within the organization. Phishing emails are also commonly used to deliver malware onto a user's system which then harvests confidential information and automates the attack process from within the network.

Research by Verizon (2015, 2016, 2017), Fire Eye (2015, 2017) and Mandiant (2004, 2010), on recent cases of the Advanced Persistent Threat (APT) involving crimeware and cyber-espionage, show that a common technique of compromising organizations is by delivering phishing emails to targeted individuals. This phishing technique of crafting attacks to fit targeted individuals is called spear phishing. The spear phishing email is often crafted to be relevant to the recipient and also appears to come from a legitimate sender, such as a colleague or company executive, often through the use of forged e-mail addresses.

Cases of phishing attacks are still on the rise despite a long history of phishing campaigns dating back to 1995 (James, 2005). The Anti-Phishing Working Group report (APWG, 2017) reported in the fourth quarter of 2016 an increase of 65% in the number of phishing attacks compared to those reported in 2015. In addition, a trend analysis of phishing attacks since 2004 show a 5,753% increase over a 12-year period. The previous report for the first quarter of 2016 (APWG, 2016) showed a 250% increase in the number of unique phishing websites since the last quarter of 2015. PhishTank, another organization that monitors cases of phishing, reported 4.5 million phishing sites in October 2016, 42,788 of which were confirmed to be active phishing sites (PhishTank, 2016).

Research by Cyveillance (2015) on the cost of phishing shows that phishing attacks are estimated to result in losses of 5.9 billion US dollars annually. News in August 2016 (Barth, 2016; BBC News, 2016) highlighted a criminal network led by a 40 year old Nigerian man called "Mike" that had scammed individuals and companies off 60 million US dollars through email scams and phishing malware. Previous research done by Hernandez, Regalado, & Villeneuve (2015) on Nigerian scammers show consistent use of email-based social engineering to defraud businesses of millions of dollars.

Investigative reports on allegations of Russia's involvement in the 2016 elections in the United States of America also show compromise through spear-phishing emails targeted at key staff in the Democratic Party (Fire Eye, 2017). In addition, there was the 2016 attack by the group Anonymous against Kenya's

Ministry of Foreign affairs. In April 2016 Anonymous posted 1 Tera Byte (TB) of sensitive data from the ministry on the dark web. After the disclosure of the breach, Kenya's ICT Cabinet Secretary explained that the hackers succeeded in gaining access to the ministry's data through phishing. An email circulated by the head of IT dated 4th August 2015 (months before the attack) tried to alert staff on the phishing attempts being sent by people impersonating the ICT administrator (Cimpanu, 2016; Obulutsa, 2016; Waqas, 2016).

Research Problem

These recent cases of phishing attacks demonstrate that it is still an active threat to users and a growing concern for organizations today. Many organizations have focused on the use of technology without giving much attention to addressing the human factor (Luo, Brody, Seazzu, & Burd, 2011). Kevin Mitnick, one of the most renowned hackers of our time, has confessed that hackers use social engineering to exploit people since they are the weakest link even in the most secure systems (Mitnick & Simon, 2002). Mitnick points out that organizations spend a lot of money developing and implementing the best security without addressing the weak human factor in the security chain.

Organizations need a credible methodology to regularly assess their susceptibility to phishing (Dodge, Carver, & Ferguson, 2007; Ferguson, 2005; Jackson, Ferguson, & Cobb, 2005; Kumaraguru, Sheng, Acquisti, Cranor, & Hong, 2007). Results of such assessments can guide the selection and implementation of appropriate countermeasures.

This study is a response to this gap. It seeks to identify a credible methodology that information security researchers can use to assess organizational exposure to phishing threats targeted at insiders. The assessment can be done regularly and a security baseline metric can be established to routinely compare with. Assessment results can be tracked over a period of time and the effectiveness of implemented countermeasures examined to see their effectiveness in reducing insider susceptibility to social engineering attacks.

Research Question

The question that this study seeks to answer is: "How can information security researchers credibly assess the vulnerability of insiders to phishing threats?"

This study aims to present a study methodology credible in the assessment of vulnerability to phishing threats. This methodology is then employed at a university in Kenya to study its vulnerability to phishing threats. The instruments used to carry out the assessment are outlined in detail and the lessons learnt in this process are then synthesized and presented in a way that can guide practice.

LITERATURE REVIEW

A review of literature reveals three key techniques that have been used to assess vulnerability to phishing threats. One of the techniques used has been the administration of phishing knowledge tests (also known as phishing IQ tests) in questionnaires or survey instruments. Wang et al. (2012) and Vishwanath et al. (2011) used questionnaires containing images of a phishing attack that had previously been launched against a university population. They asked the respondents to indicate their likelihood to respond to the phishing email that was presented. They were not able to examine actual user responses to the phishing attack because they did not stage the attack themselves. They relied on participant self-evaluation responses to gauge phishing susceptibility. Similarly, Downs et al. (2007) presented an online

questionnaire survey to 232 respondents showing images of five emails and asking them to indicate how they would respond to each email in order to gauge their phishing susceptibility. They also administered a knowledge test to gauge the participants' understanding of padlock icons and selected terminology as relates to phishing. In another related study, Tsow & Jakobsson (2007) administered an online questionnaire survey to 435 participants displaying six emails and six webpage screenshots. They asked the participants to score each screenshot, on a scale of 1-5, on how much they believed the messages to be a phishing ploy or to be genuine communication. They then used the data to evaluate both trust and deceptive tactics used in phishing scams.

A second technique commonly used in the assessment of phishing susceptibility is the conducting of lab experiments. Sheng et al. (2010), Kumaraguru, Rhee, Sheng, et al. (2007), Kumaraguru, Rhee, Acquisti, et al. (2007), Kumaraguru, Sheng, et al. (2007), Jakobsson, (2007), Jakobsson et al. (2007) and Downs et al. (2006) recruited non-expert volunteers to take part in lab experiments. Their studies involved role-play exercises staged in a lab environment where participants were asked to process a set of emails with hyperlinks. The participants were required to speak out their thought process so that the researchers could listen and identify the criteria used to distinguish phishing emails from legitimate ones. The researchers analyzed the participant feedback using what they termed as the 'think-aloud' protocol. They were then able to outline the various techniques and criteria participants used to identify phishing emails. These studies also incorporated anti-phishing training in different variations to assess the efficacy of different training approaches as treatments to address phishing susceptibility. The use of embedded training was singled out as the most effective method of delivering anti-phishing training. Another related study by Egelman et al. (2008) used a lab experiment to present 60 participants with phishing messages and to observe their interaction with browser-based warnings. The participants were divided into four groups whereby three of the groups received browser warnings when they interacted with phishing links and the control group did not. They found that 97% of the participants were susceptible to at least one phishing attack.

A third technique used in determining the susceptibility to phishing threats is the staging of real-world phishing attacks. Luo et al. (2013) asked graduate students in an information assurance class to conduct a phishing attack targeting 105 staff and faculty in the School of Management at a southwest US university. The phishing attack was designed to imitate urgent school email communications. The pretext scenario used a survey regarding possible budget cuts affecting the targeted academic and administrative staff. A total of 38 users (36%) clicked the link and 16 (15%) disclosed their usernames and password credentials on the phishing forms. Similarly, Bakhshi et al. (2009) staged a phishing attack targeting a single department in an organization with over 2,000 users. A phishing email was sent to 152 staff requesting them to install an important software update by clicking a hyperlink to an external website. A total of 35 people (23%) clicked the hyperlink and also clicked a button marked 'Proceed' to install the software update. In other related studies, Kumaraguru et al. (2009) and Kumaraguru et al. (2008) staged phishing attacks targeting a university population and a large corporation respectively. They purchased domains, set up real websites and delivered phishing emails to targeted participants. They found that 90% of participants who are vulnerable to phishing will click links within 8 hours of a phishing attack being delivered to them. They also designed their experiments to assess the effectiveness of various training options in reducing susceptibility to phishing attacks. In another study, Jagatic et al. (2007) crawled a popular social media site to extract profile information that was later used to customize phishing attacks. Results revealed a 72% success rate when the information was used to customize phishing attacks. Other studies by Dodge et al. (2007) and Jackson, Ferguson, & Cobb (2005) involved staging phishing attacks targeted at students studying at a Military Academy in West Point, Untied States. Four phishing scenarios were designed into the attack namely: clicking of hyperlinks, opening of

attachments, submitting sensitive information to online forms and the installation of downloaded applications. Results showed that 29% of those targeted clicked phishing links, 47% opened phishing attachments and 45% submitted sensitive information on a staged phishing website. The researchers were unable to stage the download and installation of a questionable application due to security and privacy concerns.

Each of these techniques has its own set of strengths and limitations as discussed hereafter. The use of phishing knowledge tests has a number of strengths. The technique is easily incorporated into questionnaires and survey instruments that can be distributed to many participants to collect data using uniform measures and verified scales. This makes it a very cost effective method of assessing phishing susceptibility and does not require the setup of technical infrastructure for phishing. The questionnaires can also be used to measure non-observable constructs associated with phishing susceptibility such as; intentions, attitudes and perceptions. The use of knowledge tests is associated with notable limitations making them unsuitable. Anandpara, Dingman, Jakobsson, Liu, & Roinestad (2007) demonstrated that these tests do not measure capabilities and skills in detecting phishing attacks. In fact, scoring highly in the tests may give participants a false sense of confidence that they are not susceptible to the threat. Additionally, knowledge tests require participants to self-report and may have elements requiring participants to recall their actions or thought processes from events that took place in the past. This can introduce measurement bias because people are known to assess themselves more favorably than they would act in practice. In addition, people tend to forget and may make up facts to fill in gaps in their recollection of past events. Participants may also respond in ways that are considered 'acceptable' to the researchers because they know they are under study. The Hawthorne effect (Parsons, 1974) explains that study participants are known to alter their behaviour due to the awareness of being studied. This leads to contamination of results. Additionally, these knowledge tests often use static content (such as screenshots) to illicit participant evaluations. Such static content is devoid of many interactive security indicators that would be available to users in real-life settings to identify phishing attacks.

The use of lab experiments in phishing assessments has its strengths. Researchers have been known to use 'think-aloud' protocols and observation techniques that provide rich insights in the participants' thought and decision-making processes. The data collected through such protocols allows for quantitative and qualitative analyses.

Conversely, lab experiments require considerable technical expertise to simulate real-world phishing attacks in a lab setting. In addition, resource constraints of a lab setup can make it difficult to engage many study participants at the same time. Past studies by Kumaraguru, Rhee, Sheng, et al. (2007), Kumaraguru, Rhee, Acquisti, et al. (2007), Kumaraguru, Sheng, et al. (2007) engaged totals of 49, 30 and 28 participants respectively, with each participant being interactively engaged during data collection. Richer engagement in study protocols could also mean more time and effort during data collection. It can also be argued that simulated environments are not comparable to real attacks. They may create a false sense of security in participants because they are not exposed the real consequences of a phishing attack. Consequently, participants may be more willing to take actions that they would otherwise not take when under a real attack as was observed by Downs et al. (2007). Furthermore, participants know they are being studied and in many cases they are primed to look out for the threat (Dhamija, Tygar, & Hearst, 2006). This heightens their awareness and alters their behaviour contrary to what would have been the case in their normal day-to-day activities. This behaviour modification contaminates the results of the study and compromises the validity and reliability of results. In addition, the selection of participants for lab studies may also introduce bias. Such recruitment often requires participants to volunteer to take part in the study and may also use convenience samples. Consequently,

there could be unique characteristics about the type of participants who take part - meaning they are not a good representation of the general population as noted by Kumaraguru, Sheng, et al. (2007) and Downs et al. (2007). This threatens the ecological validity of the study and makes it harder to generalize the findings to real-life settings and to more diverse populations.

The use of a naturalistic studies incorporating staged attacks that mimic real-world threats is arguably the recommended method of assessing susceptibility to phishing threats. Finn & Jakobsson (2007) argue that they are more effective than lab studies or knowledge tests. This is because naturalistic studies seek to observe actual behaviour in its normal context. The insiders are not made aware of the ongoing study and are expected to operate as they normally would in the absence of the study. This protects against the Hawthorne effect. In addition, Huber et al. (2009), Kumaraguru et al. (2009) and Workman (2007, 2008a) point out that such naturalistic studies have high ecological validity. Brewer & Crano (2014) explain that ecological validity is associated with studies whose settings approximate the real-world scenarios and what is everyday life for the wider population. High ecological validity enables results to be generalized to wider populations with similar real-world settings. In addition, the infrastructure required to stage phishing attacks is now readily accessible and fairly easy to setup as demonstrated by graduate students in the study by Luo et al. (2013). Researchers can purchase domains, setup web servers and carry out mass mailing to target large populations in a straightforward manner. The phishing instruments can also include active scripts and backend tools to collect a diverse collection of data about targeted users' online behaviours, even without alerting them. This avails rich data to researchers that allows them to build holistic user profiles. In addition, this data can be collected from many participants simultaneously.

Despite these advantages of using naturalistic field studies, Huber et al. (2009) and Kumaraguru et al. (2009) acknowledge that they are more difficult to conduct. It is difficult to get organizations willing to cooperate with the researcher to stage attacks that are as realistic, convincing and deceptive as would real attacks. In addition, such studies require approvals from research and ethical review boards which may be hard to get due to associated research risks. Many ethical review boards may be concerned by the use of deception and waiver of informed consent by participants (Finn & Jakobsson, 2007). Therefore, key to the success of such research is to identify an organization that is willing to have a naturalistic study conducted. Such an organization would give a site approval for the research on behalf of its population, with adequate oversight to ensure that there is no actual harm. Another challenge in delivering naturalistic studies is the technical expertise needed to deliver very realistic phishing attacks. The process often involves registration of domains, setting up of webservers, backend databases and phishing accounts.

Table 1 outlines a summary of the different techniques used to assess susceptibility to phishing.

Table 1: Critique of assessment techniques

Technique	Previous Studies Employing Technique	Strengths of Technique	Limitations of Technique
Knowledge / Phishing IQ Tests	• Wang et al. (2012) • Vishwanath et al. (2011) • Downs et al. (2007) • Tsow & Jakobsson (2007)	• Easy to administer in form of questionnaires or surveys. • Data can be collected in a uniform way using well-defined measures. • Cost effective method of collecting data from many participants. • Useful in measuring non-observable constructs, such as, intentions, attitudes and perceptions.	• Requires respondent to remember their past behaviour. Their memory may fail them. • Data collected subjected to self-reporting bias • People may assess themselves more (or even less) favorably than would actually act. • Results may be contaminated by Hawthorne-effects. • Not a true reflection of real-world attacks. • Many interactive features/tools not available to users for use in identifying phishing attacks.
Lab	• Sheng et al. (2010)	• Researchers can engage	• Researchers may find it hard to engage many

Technique	Previous Studies Employing Technique	Strengths of Technique	Limitations of Technique
Experiments	• Kumaraguru, Rhee, Sheng, et al. (2007) • Kumaraguru, Rhee, Acquisti, et al. (2007) • Kumaraguru, Sheng, et al. (2007) • Jakobsson, (2007) • Jakobsson et al. (2007) • Downs et al. (2006) • Egelman et al. (2008)	participants in "think-aloud" protocol technique to better understand their thought processes and behaviour. • Can allow richer data set involving qualitative and quantitative elements during study	participants at the same time. • Involves more time and effort in data collection. • Technical expertise is need to simulate a research environment that provides a set of features/tools to match real-world settings. • Not a true reflection of real-world attacks. • Participants are shielded from 'real' consequences. • Susceptible to Hawthorne-effects. • Results may not be generalizable.
Naturalistic Experiments	• Luo et al. (2013) • Bakhshi et al. (2009) • Kumaraguru et al. (2009) • Kumaraguru et al. (2008) • Jagatic et al. (2007) • Dodge et al. (2007) • Jackson, Ferguson, & Cobb (2005)	• Can directly and reliably observe the responses/behaviour. • High ecological validity. • Results are highly generalizable. • Fairly easy to stage. • Can target large populations. • Rich data can be collected using backend tools and scripts.	• It is difficult to get organizations willing to approve the staging of phishing attacks. • It is difficult to obtain research approvals and informed consent from participants. • Care has to be taken to ensure participants are not exposed to actual harm. • Technical expertise needed to deliver realistic/believable phishing attacks.

METHODOLOGY

This research used a naturalistic field study experiment to stage a phishing attack targeting a university population. This methodology is argued, with reasons summarized in Table 1, to be the most effective methodology when compared to the use of phishing knowledge IQ tests or lab experiments.

Research Setting

Previous researchers have found it very difficult to obtain cooperation to study information security threats in organizations (Bakhshi et al., 2009; Finn & Jakobsson, 2007; Huber et al., 2009; Vishwanath et al., 2011; Wang et al., 2012). This is a source of frustration for many information security researchers because many organizations either decline to have the study conducted altogether or restrict the publication of results (Kumaraguru et al., 2008). Some researchers opt not to conduct some elements of their study in order to obtain research approvals (Huber et al., 2009). In other cases, the research is prematurely terminated thereby negatively impacting data collection (Bakhshi et al., 2009).

There could be many reasons for this reluctance. Many organizations are wary of opening their doors for research due to the sensitivity of their systems and the confidential nature of their information and work practices (Burstein, 2008). They may not want their practices to be known to external parties, particularly competitors (they might lose intellectual property or competitive edge) or regulatory bodies (if they think their practices are deficient and may attract penalties). In addition, organizations are wary of negative publicity that may impact their bottom line due to loss of customers and revenue.

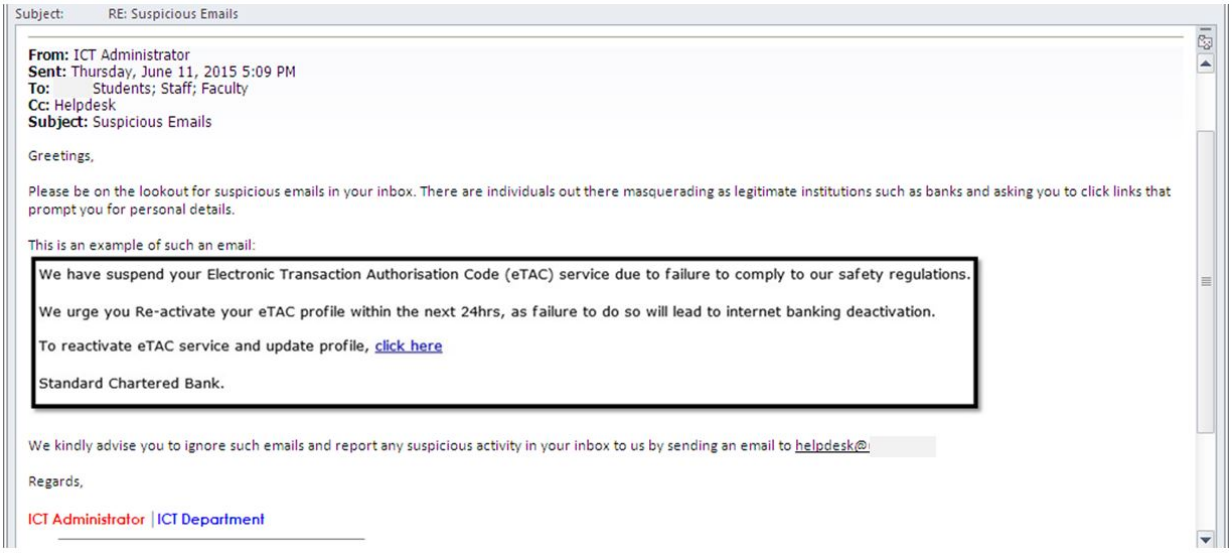
Therefore, a key criteria for selection of a research setting to study information security threats, such as phishing, is obtaining a willing organization that would give approval for conducting the research, the collection of sufficient data and publication of results (Bakhshi et al., 2009).

Getting a willing organization was an arduous task for this study. Five organizations consisting of 3 banks, 1 manufacturing company and 1 public utility company were contacted over a 14 month period to obtain approvals to conduct the research. All these institutions declined. The organization that was willing to allow this research to be conducted was a private university located in Nairobi, Kenya.

The selection of a university research site to study similar information security threats has been done in previous studies (Arachchilage & Love, 2013; Dodge et al., 2007; Finn & Jakobsson, 2007; Liang & Xue, 2010; Luo et al., 2013; Vishwanath et al., 2011; Wang et al., 2012; Workman, 2007, 2008a, 2008b). Universities are a suitable research site because they encourage research and the discovery of knowledge as long as the research is conducted ethically and does not harm the university community (Finn & Jakobsson, 2007).

Another key in determining the research setting was the selection of a naturalistic environment where the threat phenomenon was known to occur and could be observed without alerting study participants of the ongoing study. This required staging of attacks mimicking real-life threats and targeting study participants who were not aware of the ongoing research (Bakhshi et al., 2009; Finn & Jakobsson, 2007; Huber et al., 2009; Vishwanath et al., 2011). These staged attacks needed to be conducted in a way that made them as convincing and deceptive as real attacks.

The institution selected for this study had been a target of numerous social engineering attacks through phishing and wanted assistance in addressing the issue. Many of the attacks sought to obtain the confidential data, particularly passwords, through phishing emails as illustrated in Figure 1. Any identifying information in the figure has been greyed out to protect the identity of the institution.


Subject: RE: Suspicious Emails

From: ICT Administrator
Sent: Thursday, June 11, 2015 5:09 PM
To: Students; Staff; Faculty
Cc: Helpdesk
Subject: Suspicious Emails

Greetings,

Please be on the lookout for suspicious emails in your inbox. There are individuals out there masquerading as legitimate institutions such as banks and asking you to click links that prompt you for personal details.

This is an example of such an email:

We have suspend your Electronic Transaction Authorisation Code (eTAC) service due to failure to comply to our safety regulations.

We urge you Re-activate your eTAC profile within the next 24hrs, as failure to do so will lead to internet banking deactivation.

To reactivate eTAC service and update profile, click here

Standard Chartered Bank.

We kindly advise you to ignore such emails and report any suspicious activity in your inbox to us by sending an email to helpdesk@

Regards,

ICT Administrator | ICT Department

Figure 1: Sample phishing attack previously targeted at insiders

Other attacks sought to install malware on information systems through malicious attachments. The organization had been hit by numerous malware infections and ransomware attacks through this social engineering technique. The organization resonated with the proposed research and wanted assistance in assessing their exposure to the phishing threat.

Ethical Considerations

Research ethics relates to moral choices and decision making concerning research conduct (Greener, 2008). Various principles have to be upheld in the course of a research; these include: honesty, integrity, objectivity, respect for intellectual property, confidentiality and protection of research participants. Diener & Crandall (1978) highlight four main issues relating to research ethics: harm to participants,

deception, invasion of privacy and lack of informed consent. Finn & Jakobsson (2007) point out that there are various ethical issues to be addressed when conducting information security research particularly when staging naturalistic experiments involving deception. This research took special care to address these ethical concerns.

Institutional approval to conduct research at the university was obtained from its research office and information communication and technology (ICT) department. This provided a site approval to conduct the research and collect data from the insiders. In addition, an Institutional Research Board (IRB) gained approval of the research proposal, data collection procedures and methodologies.

These various levels of approval were necessary in order to ensure that the study did not pose any actual harm to the participants or the institution. Two senior ICT administrators were attached to the research to review the phishing instruments to ensure that none of the technical components harmed the organization's information system or collected sensitive data from the insiders. The IRB approval signified that the research was found to meet the required ethical standards and was not going to be harmful to the participants or the organization.

Finn & Jakobsson (2007) point out that the deceptive nature of naturalistic studies makes it difficult to obtain informed consent from participants. This was also true for this study. However, the site approvals and oversight granted by the research office, ICT department and IRB protected the participants from adverse effects.

Diener & Crandall (1978) differentiate confidentiality and privacy and emphasize the need for research to fulfill these two key ethical considerations. Confidentiality is upheld in all stages of the research by making sure that study participants are anonymized, and no data is personally identifiable to them. Privacy regards the usage of the research data and this study ensures that the detailed raw data is not disclosed to other entities other than the researcher and the appointed academic supervisory teams. In addition, and published results are reported in collective terms where the organization or study participants are not identifiable. This study was bound by confidentiality and privacy requirements. Therefore, participant and institution data was anonymized and reported in collective terms.

Phishing Instruments

The development of the phishing instruments for this study was guided by the recommendations and lessons learnt from previous studies by Luo et al. (2013), Arachchilage & Love (2013), Vishwanath et al. (2011) and Bakhshi et al. (2009).

First, typical samples of phishing attacks launched against the insiders in the organization were studied. The ICT administrators attached to the research provided 12 samples of recent phishing attacks that had been targeted at the organization's insiders. Characteristics that made the phishing attacks successful were identified in collaboration with the ICT administrators. The attacks that closely imitated the organization's communication techniques and the look and feel were seen to be most deceptive. Therefore, the phishing instruments were designed to closely conform to the layout, fonts, look and feel used within the organization.

Secondly, a domain that imitates the organization's domain was selected. Instead of using the registered domain ending with "ac.ke," the researcher registered a domain that ended with "or.ke." The email address "helpdesk@universityX.or.ke" was used and the website was hosted on "universityX.or.ke." This ensured that the email and website address used to conduct the attack would closely imitate the

organization's legitimate addresses while allowing for knowledgeable insiders to identify the attack by picking up an inconsistency in the addressing. This strategy is advocated by Luo et al. (2013).

The next step in the process involved the selection of a pretext scenario that would be perceived as a natural event. The scenario would then guide the development of content for the phishing email and message. The guidelines by Luo et al. (2013) and Vishwanath et al. (2011) were used to guide the design of the pretext scenario. A topic that was current and relevant to the organization was selected. The organization had a limited capacity email server and consequently users were only allowed 2GB of email space. This meant that users regularly received 'mailbox full' notifications indicating they had exhausted their allocated quota. The pretext scenario took advantage of this and advertised an opportunity for the users to increase their allocated email quota. Time pressure was also put on the users to respond urgently in order to prevent discontinuation of service similar to the Luo et al. (2013) study.

A data collection website developed in HTML5, CSS and PHP with a MySQL database was hosted on the registered domain and tested to ensure it ran without errors. In addition, the ICT administrators attached to the study reviewed the code and backend database to ensure that no malware was delivered and no sensitive or confidential data was collected and stored. This protected the insiders from actual harm as was required by directives from the research office and Institutional Review Board.

Figure 2 depicts the login page of the website. Appendix I provides the source code and Appendix II provides the Cascading Style Sheets (CSS) for the page. Any identifying information has been removed from the content to protect the identity of the institution.

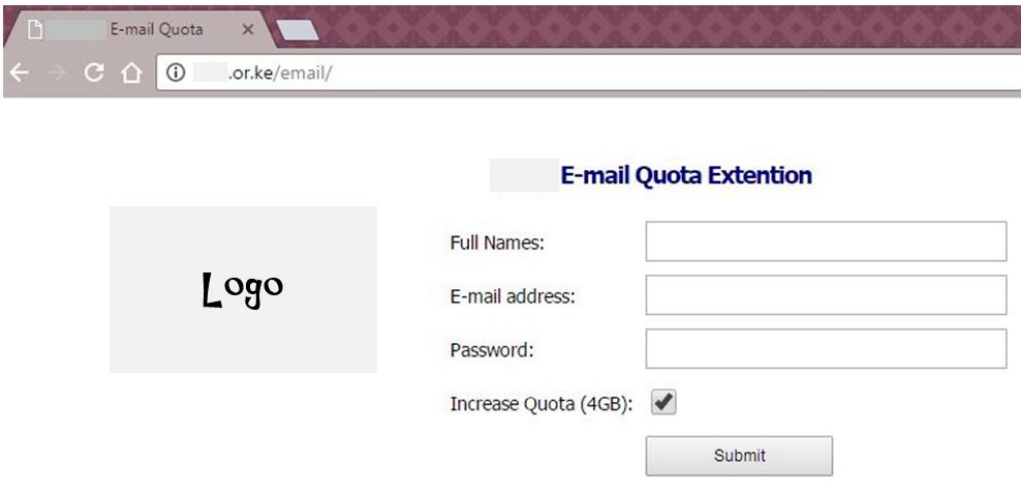


Figure 2: Phishing Website Login

Next, targeted phishing emails were sent to selected insiders. The emails were staged as spear phishing emails using the first name and surname to personalize the message. The message seemed to have been sent from the institution's helpdesk by an ICT administrator. This imitated the means of communication commonly used by the institution when sending IT related information to the users.

The email had the 'look' and 'feel' of the usual email messages from ICT administrators. It was carefully composed not to have spelling mistakes or sloppy content so that recipients do not superficially dismiss it. The variable fields in the email were filled in using mail merge. These fields were: first name, last name and email address. Figure 3 shows the mail merge template setup using the mail merge feature

on Microsoft Office Word 2013. Any identifying information has been greyed out to protect the identity of the institution.

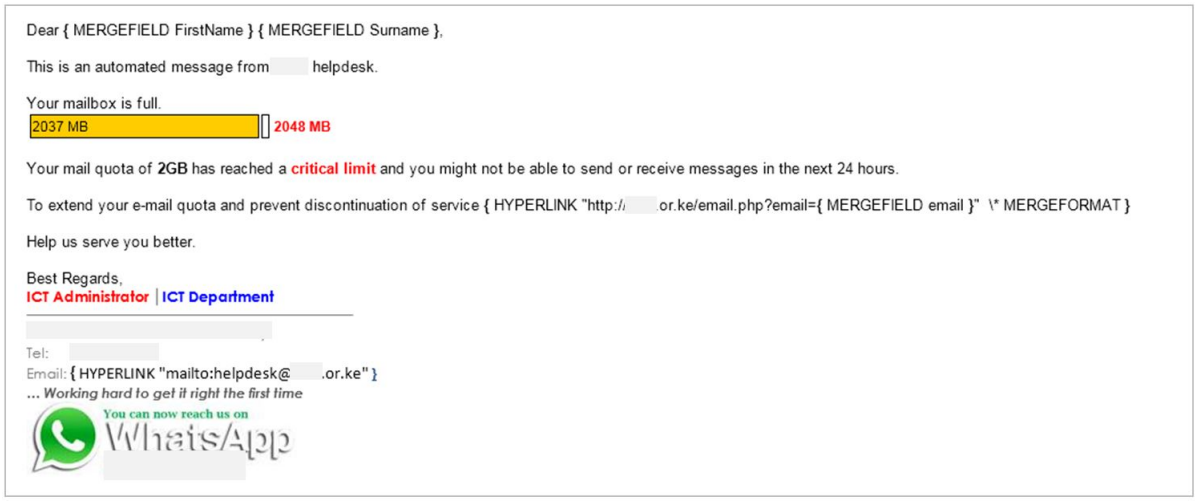


Figure 3: Phishing Mail Merge Template

The administration of emails was automated using mail merge working together with Microsoft Outlook 2013. Figure 4 shows the resulting phishing email that was sent to a sample of targeted insiders. Please note that identifying information has been greyed out to protect the identity of the institution.

Figure 4: Phishing Email

These phishing instruments collected various data items for study. The phishing email tracked when the email was successfully delivered and opened. In addition, the phishing email had a hyperlink in which the words "click here" were highlighted in blue and underlined. This hyperlink did two things. First, it directed the person to the phishing website by opening their default browser and loading the phishing website's address. Secondly, it passed on a unique identifier as a pre-filled parameter to the landing page. This means it was possible to distinctively track all the people who visited the website.

The phishing website ran active scripts that recorded a timestamp of when the page was loaded, the identifier registered from the forwarding email and various parameters about the system accessing the page including the IP address, browser and Operating System. The source code of the background script is provided in Appendix III. This means that even if the user did not interact further with the website, just loading it gave a lot of valuable information.

The other way data was collected was when a person filled in the form on the website. This involved submitting the following details: full name, email address and password. The email address was already pre-filled if the person clicked the hyperlink from the phishing email. This communicated some level of sophistication to users that was designed to make the website more trustworthy. When a person filled in the form and clicked the submit button their password was neither captured nor transmitted as a design requirement. This prevented the capturing of confidential information and protected the institution from actual harm. The webpage also had error validation to ensure that the submit functionality did not work if the required form fields were blank.

Sampling

In the context of this study, the effective population was all the insiders who had active email accounts on the university's system. These were all the potential targets of any phishing attack directed at the university using its domain. The domain account management system was queried by its system administrator to provide the exact number of insiders at the time of the study.

Strata	Number
Students	7,729
Staff	312
Adjunct Faculty	158
Full-time Faculty	141
Management	13
Interns	9
Mailing List Users	7
Unknown	36
Total Insiders	**8,405**

Table 2: Sampling Frame

The university campus had a total of 8,405 insiders active on its information system. Of these, 7,729 were students, 312 were staff members, 158 were adjunct faculty, 141 were full-time faculty, 13 were management, 9 were interns, 7 accounts were mailing list accounts and 36 could not be classified in any of these categories due to insufficient metadata. Table 2 illustrates this sampling frame.

This study employed a probability sampling technique to allow the results to be generalizable to the population. Bhattacherjee (2012) explains that in probability sampling, each entity in the population has a non-zero chance of being selected in the sample. In addition, random selection techniques are

employed in the sampling process. This ensures that sample statistics are unbiased estimates of what is in the population.

The specific technique selected was proportional stratified random sampling. The process as outlined by Bhattacherjee (2012) involves dividing the sampling frame into non-overlapping groups called strata. Thereafter a simple random sample is drawn from each stratum in what is called multi-stage random sampling. This ensures that the strata with few members is not oversampled and the resulting sample has similar ratios for the different strata.

The determination of sample size used the Cochran (1977) formula. It targeted a 95% confidence level and a very low margin of error at 1%. The proportion of sampling in the population was set at 50% to give maximum variability. This resulted in a sample size of 4,483 being extracted from the population of 8,405 insiders. To prevent under-sampling or over-sampling per strata, proportional stratified random sampling was done to determine the actual composition of the sample per strata. The numbers per strata selected for the sample are represented in Table 3.

Strata	Number	Proportion	Size in Sample
Students	7,729	91.96%	4,122
Staff	312	3.71%	166
Adjunct Faculty	158	1.88%	84
Full-time Faculty	141	1.68%	75
Management	13	0.15%	6
Interns	9	0.11%	4
Mailing List Users	7	0.08%	7
Unknown	36	0.43%	19
Total	**8,405**	**100%**	**4483**

Table 3: Sample Size

The size in sample for each stratum was then chosen using simple random sampling with the aid of a random number generator. To do this, the dataset associated with the 8,405 users were loaded onto a Microsoft Excel 2013 workbook. Each row of the workbook was associated with one user. The entries were grouped sequentially according to the strata outlined in Table 3. Next, a new column was added on the workbook to contain the random number. The random number was generated using the RAND() function entered as a formula =RAND() for every cell in the column. This ensured that each user entry was assigned a random number. After the random numbers were assigned, the entries were sorted in ascending order while still maintaining the strata groupings. Finally, the required size in sample, say 'n_s', was selected by choosing the first n_s entries in each stratum. These entries were transferred to a new workbook representing the selected sample dataset of 4,483 users.

RESULTS AND ANALYSIS

The phishing experiment ran for 40 days. It had to be stopped because a prominent social media activist and blogger, who was also a student at the university, called for the phishing to be investigated and stopped. His comment was posted on the university's social media and within a few hours had reached many people within the university. The social media post is illustrated in Figure 5. Any identifying information has been greyed out to protect the identity of the institution.

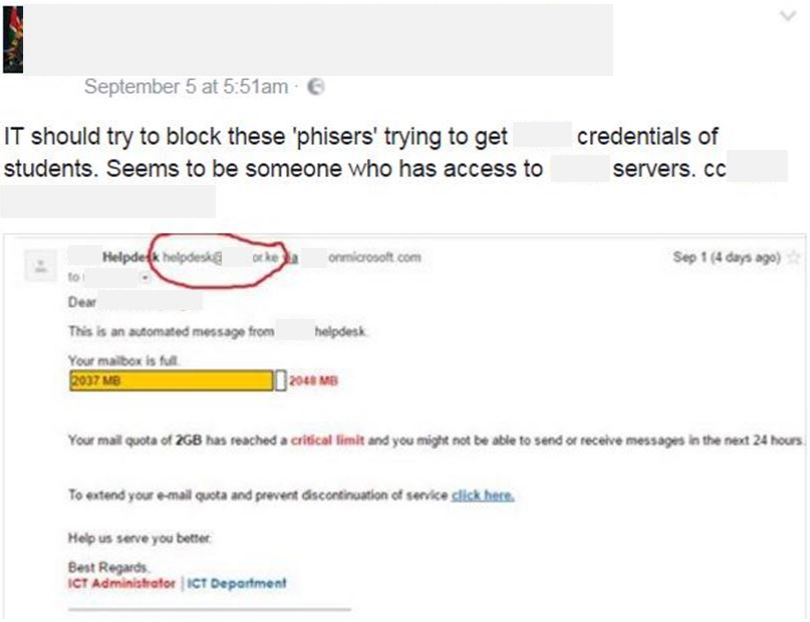


Figure 5: Phishing alert sent on social media

This prompted the administration at the university to call off the exercise due to the alarm raised. The ICT director, who had been involved in the research approvals and was aware of its progress, instructed his team to send out alerts to the entire university community informing them of the nature of the research and allaying any concerns of an actual threat. This demonstrates the power informed and vocal insiders have in identifying threats and alerting their communities to curtail targeted attacks.

By this time, all the 4,483 insiders who were sampled from the university community had already been sent phishing emails through their official university accounts. The email system returned delivery failures for 138 of the emails indicating that there was a problem with these email accounts. This meant that 4,345 phishing emails were delivered to the insiders' official email accounts. Statistics on interaction with the phishing email were low. There was no response or interaction with the phishing email by 4,104 of the targeted sample.

Category	Number	Percentage
Sample size targeted with phishing email	4,483	**100%**
E-mail delivery failures	138	**3.08%**
Did not read/interact with phishing email	4,104	**91.54%**
Insiders that participated	**241**	**5.37%**

Table 4: Response Rate Statistics

The number of insiders who participated in the experiment were 241. This was 5.37% of the total number sampled. These are the people who received the phishing emails and opened them. Read receipts were setup in Microsoft Outlook to give this indication. Data collected on the backend database indicated a total of 98 clicks on the phishing hyperlink. These clicks were associated with 74 unique insiders since some clicked the phishing hyperlink multiple times, as indicated by repeated entries in the backend database. In addition, the form on the phishing website was filled in 72 times with 65 form-fills being unique and the others being repeated entries. This shows 87.84% of the insiders who were

susceptible to phishing emails went ahead to disclose passwords that would enable an attacker gain access to the organization's systems.

The response rate per strata is provided in Table 5. Results shows that interns (25%), staff (22.89%), full-time faculty (17.33%) and mailing list users (14.29%) had higher response rates in proportion to the numbers targeted per strata. Students had a very low percentage (0.49%) of successfully phished per strata despite having the highest number in sample.

Strata	Size in Sample	Successfully Phished	Proportion
Students	4,122	20	0.49%
Staff	166	38	22.89%
Adjunct Faculty	84	1	1.19%
Full-time Faculty	75	13	17.33%
Management	6	0	0%
Interns	4	1	25%
Mailing List Users	7	1	14.29%
Unknown	19	0	0%
Total	**4483**	**74**	**100%**

Table 5: Response Rate per Strata

DISCUSSION AND RECOMMENDATIONS

The study presents interesting findings and valuable lessons. These lessons are hereafter synthesized into guidelines for practice. The first guideline advocates for the use of naturalistic field experiments when assessing how susceptible insiders are to phishing attacks. This is because naturalistic studies allow a direct observation of actual behaviour and this provides a more reliable assessment of phishing susceptibility than self-reported measures. The insiders being studied are not alerted about the assessment and are expected to act as they normally would in their real-world settings. This provides a high ecological validity of results and makes them generalizable to the population. In addition, rich data informing researchers on user behaviours can be collected from large populations with relative ease.

The second guideline emphasizes the need to obtain research approvals from the organization where the study is to be conducted. Getting permission to conduct information security research is often an arduous task, as was with this study. This however does not preclude the need for research approvals. It is important to get site approvals from the necessary representatives of the institution where the research will be conducted. It is important to obtain an ethical review approval from an IRB to ensure that the research protocol protects participants from actual harm. In this study, approvals were obtained from the university's research office, ICT department and IRB. These layers of review protected the institution and its insiders from adverse effects during the staged phishing attack.

The third guideline relates to the development and setup of phishing instruments. In this study, phishing was conducted using targeted spear phishing emails and also by setting up a phishing website. Before any phishing instrument was developed care was taken to design them to be convincing. Previous phishing attacks targeted at the institution were studied and the characteristics of regular communication were noted. In addition, a pretext scenario that was relevant to the current affairs at the organization was chosen. These considerations during design ensure that the staged attack is not easily dismissed without eliciting interaction from those targeted. In addition, a phishing domain that was deceptively similar to

the organization's operational domain was registered. Organizations should closely monitor domains that are very similar to their operational domains. These could be deceptive variants of their operational domain but also those ending with different suffixes such as .com, .org or even country suffices such as .or.ke, as was used in this study. The information technology or security teams at the organization should probably go a step further to buy such domains instead of leaving them available for outsiders to acquire. It is not a very expensive venture since registering a domain could cost as low as 10 dollars per year, as was the case in this study. This study also setup a phishing email address to imitate the ICT helpdesk correspondence to deliver spear phishing emails. The phishing emails and website used in this study were designed with active scripts that did a lot of background work. This highlights a very important point. Phishing is not considered successful only when a person fills in sensitive information on a web form. Attackers collect valuable information right from the time a person opens their emails or loads their websites. Current phishing scams are very sophisticated. Emails and webpages contain a lot of active scripts that harvest information from user systems and even install malware without visibly alerting users. A key contribution of this research involves presenting the actual code that was used to implement the phishing webpages and login forms, the active background scripts that harvested system details and also the mail merge templates that were used to deliver phishing emails.

The fourth guideline relates to population sampling. Care should be taken to ensure a representative sample is drawn from the population before the staged phishing attack is delivered. This study advocates for a probability sampling technique because it allows the results of the assessment to be generalizable to the wider population under investigation. The study outlines the sampling process in detail. It starts by extracting a dataset of user accounts from the information system and arranging these accounts according to the different functional divisions in the organization. This ensures that no division or functional group is over-sampled or under-sampled. Thereafter, actual phishing targets are selected randomly and in proportion to the strata representation in the population. This rigorous sampling process ensures the sample used in the assessment is a good representation of the population and this in turn allows the results and lessons learnt to be generalized to the wider organization without bias.

The fifth guideline relates to the actual execution of the staged phishing attack. In this study, the staged attack ran for 40 days. It is important to allow a similar amount of time for the exercise to run to account for differences in email responsiveness among the population. Some participants may respond immediately while others may defer their email processing for days or even weeks. The days allocated to the exercise should also factor times where participants are expected to be on holiday or may have limited access to their accounts. During the execution it would be important to look out for any insiders who alert the community about the attack and to examine the reaction of the organization to this alert. These vocal insiders could affect the event of an actual attack and could present a very important countermeasure that organizations should focus on when addressing attacks. In this study, a prominent blogger was able to raise an alarm and rally action through social media. Within a few hours, an alert of the ongoing threat had been circulated throughout the entire institution. Study protocols should collect data and assess the effectiveness of such countermeasures in curtailing attacks. Organizations should invest in channels through which users can quickly report suspected attacks and through which information can be shared with the wider population to frustrate the efforts of attackers. They should turn each user on their system into an intrusion detection agent with the skill and capability to detect threats and to sound an alarm for action.

The sixth guideline relates to the actual results of the study. This study suffered a significant attrition with regards to the actual users who engaged with the phishing emails. Only 5.37% of the 4,483 sampled participants engaged with the phishing email. This is comparable to the study by Mohebzada, El Zarka,

BHojani, & Darwish (2012) where over 10,000 phishing emails were sent to faculty, staff, students and alumni of a university. Two types of phishing emails were sent. The first phishing email had an 8.74% success rate and the second 2.05%. Low phishing rates could be an indication that the participants identified the phishing scam and chose not to engage with the phishing email or website. Although the phishing rates were low, it only takes a few users to compromise an information system. Once an attacker is successful with some systems, these can then be used as a pivot point to work into the rest of the organization (Ali, 2015).

Some lessons can be learnt from our study to help future studies increase the number of users who interact with the phishing instruments. Firstly, it would be important to confirm that the emails that will be targeted are operational and that no delivery failures will be experienced. Secondly, it is important to confirm whether users regularly engage with their emails. Discussions with the ICT staff attached to the study revealed that it could be that few people used their official university accounts for correspondence. Students (who were the largest number in the sample) had an option of using alternative email addresses to receive communication from the university. This meant that they had no imperative to use their official email accounts. Instead they preferred to use private email accounts mainly from Google, Hotmail or Yahoo. If official institution email addresses are not used regularly, then personal emails registered for official communication should also be included in the sample. Thirdly, increasing the study period would also give users a longer time to review their emails and thereby possibly increase their participation. Fourthly the assessment could also target other channels to deliver the phishing attack, for example, using organizational social media accounts and telephone chat groups.

CONCLUSION

Phishing is still a very prevalent form of social engineering attack leveraged against organizations today. Recent reports have shown that it is a common method of compromising organizations and spreading Advanced Persistent Threats (APTs). It is important that organizations take steps to assess their level of risk and exposure to this attack. This study presents a way in which organizations can use a naturalistic study to objectively assess their exposure to phishing threats. Organizations can use such naturalistic experiments to regularly determine the extent to which their users can succumb to phishing attacks. The data collection instruments used in the naturalistic field study are not difficult to assemble. This study makes an important contribution by outlining the actual tools used to stage the phishing attack in detail. Such assessments can be run on a routine basis to provide a security baseline metric from which to compare from time to time. The results of the assessments can be very useful in designing countermeasures, one of which is discussed in this study. Insiders can be equipped to detect attacks and channels to alert the wider community can be provided to them. This would inevitably provide an essential component of strengthening the overall information security of an organization, particularly from a people-perspective, which organizations often leave unaddressed.

ACKNOWLEDGMENTS

This work was funded through the German Academic Exchange Service (DAAD) In-Country/ In-Region Scholarships for Postgraduates.

REFERENCES

Ali, A. (2015). Social Engineering: Phishing latest and future techniques. Retrieved March 10, 2016 from https://www.researchgate.net/publication/274194484

Anandpara, V., Dingman, A., Jakobsson, M., Liu, D., & Roinestad, H. (2007). Phishing IQ Tests Measure Fear, Not Ability (pp. 362–366). Presented at the International Conference on Financial Cryptography and Data Security, Trinidad and Tobago: Springer Berlin Heidelberg.

APWG, A.-P. W. G. (2016). Phishing Activity Trends Report: 1st Quarter 2016. Anti-Phishing Working Group. Retrieved March 9, 2016 from https://docs.apwg.org/reports/apwg_trends_report_q1_2016.pdf

APWG, A.-P. W. G. (2017). Phishing Activity Trends Report: 4th Quarter 2016. Anti-Phishing Working Group. Retrieved December 12, 2017 from https://docs.apwg.org/reports/apwg_trends_report_q4_2016.pdf

APWG, A.-P. W. G. (2018). Phishing Activity Trends Report: 3rd Quarter 2017. Anti-Phishing Working Group. Retrieved April 30, 2018 from http://docs.apwg.org/reports/apwg_trends_report_q3_2017.pdf

Arachchilage, N. A. G., & Love, S. (2013). A Game Design Framework for Avoiding Phishing Attacks. *Computers in Human Behavior*, *29*, 706–714.

Bakhshi, T., Papadaki, M., & Furnell, S. (2009). Social Engineering: Assessing Vulnerabilities in Practice. *Information Management & Computer Security*, *17*(1), 53–63. https://doi.org/10.1108/09685220910944768

Barth, B. (2016). Don't be like "Mike": Authorities arrest mastermind of $60M online scam operation. SC Magazine.

BBC News. (2016). Online fraud: Top Nigerian scammer arrested. BBC News. Retrieved October 3, 2016 from http://www.bbc.com/news/world-africa-36939751

Bhattacherjee, A. (2012). *Social Science Research: Principles, Methods, and Practices* (2nd ed.). University of South Florida Scholar Commons.

Brewer, M. B., & Crano, W. D. (2014). Research Design and Issues of Validity. In H. T. Reis & C. M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology* (2nd ed., pp. 3–16). New York: Cambridge University Press.

Burstein, A. J. (2008). *Toward a Culture of Cybersecurity Research* (UC Berkeley Public Law Research Paper No. 1113014). Retrieved from http://dx.doi.org/10.2139/ssrn.1113014

CERT, I. T. T. (2013). Unintentional Insider Threats: A Foundational Study. Software Engineering Institute, Carnegie Mellon University. Retrieved October 17, 2013 from http://resources.sei.cmu.edu/asset_files/TechnicalNote/2013_004_001_58748.pdf

Cimpanu, C. (2016). Anonymous Hackers Leak 1 TB of Documents from Kenya's Ministry of Foreign Affairs. Retrieved September 12, 2016, from http://news.softpedia.com/news/anonymous-hackers-leak-1tb-of-documents-from-kenya-s-ministry-of-foreign-affairs-503518.shtml

Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). New York: Wiley.

Cyveillance. (2015). The Cost of Phishing: Understanding the True Cost Dynamics Behind Phishing Attacks.

Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why Phishing Works. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 581–590). ACM.

Diener, E., & Crandall, R. (1978). *Ethics in Social and Behavioral Research*. University of Chicago Press.

Dodge, R. C., Carver, C., & Ferguson, A. J. (2007). Phishing for User Security Awareness. *Computers & Security*, *26*, 73–80.

Downs, J. S., Holbrook, M., & Cranor, L. F. (2006). Decision Strategies and Susceptibility to Phishing. In *Proceedings of the second symposium on Usable privacy and security* (pp. 79–90). ACM.

Downs, J. S., Holbrook, M., & Cranor, L. F. (2007). Behavioral Response to Phishing Risk. In *Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit* (pp. 37–44). ACM.

Egelman, S., Cranor, L. F., & Hong, J. (2008). You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM.

Ferguson, A. J. (2005). Fostering e-mail security awareness: The West Point Carronade. *Educase Quarterly*, *28*(1), 54–57.

Finn, P., & Jakobsson, M. (2007). Designing and Conducting Phishing Experiments. *IEEE Technology and Society Magazine, Special Issue on Usability and Security*, *26*(1), 46–58.

Fire Eye. (2015). APT30 And The Mechanics Of A Long-Running Cyber Espionage Operation. Retrieved August 16, 2017 from https://www.fireeye.com/current-threats/threat-intelligence-reports.html

Fire Eye. (2017). APT28: At The Center Of The Storm. Retrieved August 16, 2017 from https://www.fireeye.com/current-threats/threat-intelligence-reports.html

Greener, S. (2008). *Business Research Methods*. BookBoon.

Hernandez, E., Regalado, D., & Villeneuve, N. (2015). An Insider Look Into the World of Nigenrian Scammers. Fire Eye. Retrieved August 16, 2017 from https://www.fireeye.com/current-threats/threat-intelligence-reports.html

Huber, M., Kowalski, S., Nohlberg, M., & Tjoa, S. (2009). Towards Automating Social Engineering Using Social Networking Sites (Vol. 3, pp. 117–124). Presented at the Computational Science and Engineering, IEEE.

Jackson, J. W., Ferguson, A. J., & Cobb, M. J. (2005). Building a University-wide Automated Information Assurance Awareness Exercise. (p. T2E7-11). Presented at the 35th ASEE/IEEE Frontiers in Education Conference, Indianapolis, IN, USA: IEEE.

Jagatic, T., Johnson, N., Jakobsson, M., & Menczer, F. (2007). Social Phishing. *Communications of the ACM*, *50*(10), 94–100.

Jakobsson, M. (2007). The Human Factor in Phishing. *Privacy & Security of Consumer Information*, *7*(1), 1–19.

Jakobsson, M., Tsow, A., Shah, A., Blevis, E., & Lim, Y.-K. (2007). What Instills Trust? A Qualitative Study of Phishing. In *Financial Cryptography and Data Security* (pp. 356–361). Springer Berlin Heidelberg.

James, L. (2005). *Phishing Exposed*. Syngress.

Kumaraguru, P., Cranshaw, J., Acquisti, A., Cranor, L., Hong, J., Blair, M. A., & Pham, T. (2009). School of Phish: A Real-Word Evaluation of Anti-Phishing Training. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS)*. Mountain View, CA, USA: ACM.

Kumaraguru, P., Rhee, Y., Acquisti, A., Cranor, L. F., Hong, J., & Nunge, E. (2007). Protecting People from Phishing: The Design and Evaluation of an Embedded Training Email System. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 905–914). ACM.

Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L. F., & Hong, J. (2007). Getting Users to Pay Attention to Anti-Phishing Education: Evaluation of Retention and Transfer. In *Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit* (pp. 70–81). ACM.

Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F., & Hong, J. (2007). Teaching Johnny Not to Fall for Phish. *CM Transactions on Internet Technology (TOIT)*, *10*(2), 7.

Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L., & Hong, J. (2008). Lessons From a Real World Evaluation of Anti-Phishing Training. Presented at the eCrime Researcher's Summit, Anti-Phishing Working Group (APWG).

Liang, H., & Xue, Y. (2010). Understanding Security Behaviors in Personal Computer Usage: A Threat Avoidance Perspective. *Journal of the Association for Information Systems (JAIS)*, *11*(7), 394–413.

Luo, X. (Robert), Brody, R., Seazzu, A., & Burd, S. (2011). Social Engineering: The Neglected Human Factor for Information Security Management. *Information Resources Management Journal (IRMJ)*, *24*(3), 1–8.

Luo, X. (Robert), Zhang, W., Burd, S., & Seazzu, A. (2013). Investigating Phishing Victimization with the Heuristic-Systematic Model: A Theoretical Framework and an Exploration. *Computers & Security*, *38*, 28–38.

Mandiant. (2004). APT1: Exposing One of China's Cyber Espionage Units - FireEye. Mandiant. Retrieved October 6, 2016 from https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf

Mandiant. (2010). M-Trends: The Advanced Persistent Threat. Mandiant. Retrieved October 6, 2016 from https://www.fireeye.com/blog/threat-research/2010/01/m-trends-advanced-persistent-threat-malware.html

Mohebzada, J. G., El Zarka, A., BHojani, A. H., & Darwish, A. (2012). Phishing in a University Community: Two Large Scale Phishing Experiments (pp. 249–254). Presented at the International Conference on Innovations in Information Technology (IIT), IEEE.

Obulutsa, G. (2016). Hackers leak stolen Kenyan foreign ministry documents. Retrieved September 12, 2016 from http://www.reuters.com/article/us-cyber-kenya-idUSKCN0XP2K5

Parsons, M. H. (1974). What happened at Hawthorne? *Science*, *183*(4128), 922–932.

PhishTank. (2016). PhishTank Stats. Retrieved October 14, 2016 from https://www.phishtank.com/stats.php

Tsow, A., & Jakobsson, M. (2007). Deceit and Deception: A Large User Study of Phishing. Indiana University.

Verizon. (2015). 2015 Data Breach Investigations Report (DBIR). Retrieved July 16, 2015 from http://news.verizonenterprise.com/2015/04/2015-verizon-dbir-report-security/

Verizon. (2016). 2016 Data Breach Investigations Report (DBIR). Retrieved July 16, 2016 from http://www.verizonenterprise.com/resources/reports/rp_DBIR_2016_Report_en_xg.pdf

Verizon. (2017, April). 2017 Data Breach Investigations Report (DBIR). Retrieved May 16, 2017 from http://www.verizonenterprise.com/verizon-insights-lab/dbir/2017/

Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, R. H. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems*, *51*(3), 576–586.

Wang, J., Herath, T., Chen, R., Vishwanath, A., & Rao, R. H. (2012). Phishing Susceptibility: An Investigation Into the Processing of a Targeted Spear Phishing Email. *IEEE Transactions on Professional Communication*, *55*(4), 345–362.

Waqas. (2016). Anonymous Leaks 1TB of Data from Kenya's Ministry of Foreign Affairs. Retrieved September 12, 2016 from https://www.hackread.com/anonymous-hacks-kenya-ministry-foreign-affairs/

Workman, M. (2007). Gaining Access with Social Engineering: An Empirical Study of the Threat. *Information Systems Security*, *16*(6), 315–331.

Workman, M. (2008a). A Test of Interventions for Security Threats from Social Engineering. *Information Management & Computer Security*, *16*(5), 463–483.

Workman, M. (2008b). Wisecrackers: A Theory-Grounded Investigation of Phishing and Pretext Social Engineering Threats to Information Security. *Journal of the American Society for Information Science and Technology*, *59*(4), 662–674.

APPENDIX I: INDEX HTML PAGE SOURCE CODE

```
<?php
        session_start();

        // Initialize variables
        $username="";
        $email="";
        $passwordErr="";
        $nameErr ="";
        $emailErr ="";
        $passwordErr ="";
        $isValidUsername=0;
        $isValidEmail = 0;
        $isValidPassword = 0;

        function test_input($data) {
                $data = trim($data);
                $data = stripslashes($data);
                $data = htmlspecialchars($data);
                return $data;
        }

        if($_SERVER["REQUEST_METHOD"] == "POST") {
                // Form submitted
```

```
//-------------------------Form Validation Start---------------------//
if (empty($_POST["username"])) {
        $nameErr = "Name is required";
        $isValidUsername = 0;
} else {
        $username = test_input($_POST["username"]);
        if (!preg_match("/^[a-zA-Z ]*$/",$username)) {
                $nameErr = "Only letters and white space allowed";
                $isValidUsername = 0;
        }
        else {
                $isValidUsername = 1;
        }
}

if (empty($_POST["email"])) {
        $emailErr = "E-mail is required e.g. username@uni.ac.ke";
        $isValidEmail = 0;
} else {
        $email = test_input($_POST["email"]);
        $regex = '/^[_a-z0-9-]+(\.[_a-z0-9-]+)*@[a-z0-9-]+(\.[a-z0-9-]+)*(\.[a-z]{2,4})$/';
        if (!preg_match($regex, $email)) {
                $emailErr = "$email is not a valid email address";
                $isValidEmail = 0;
        }
        else {
                $isValidEmail = 1;
        }
}

if (empty($_POST["password"])) {
        $passwordErr = "Password is required";
        $isValidPassword = 0;
} else {
        $password = md5($_POST["password"]);
        $isValidPassword = 1;
}

//-------------------------Form Validation End---------------------//

//-------------------------Database Connection Start---------------------//
if ($isValidUsername && $isValidEmail && $isValidPassword){
        //Set up connection to database
        define('DB_SERVER', 'SERVER_NAME');
        define('DB_USERNAME', 'USER_NAME');
        define('DB_PASSWORD', 'PASSWORD');
        define('DB_DATABASE', 'DB_NAME');
        $db = mysqli_connect(DB_SERVER,DB_USERNAME,DB_PASSWORD,DB_DATABASE);

        if (!$db) {
                die("Connection failed: " . mysqli_connect_error());
        }

        //mysqli_real_escape_string used to prevent SQLi
        $username = mysqli_real_escape_string($db,$username);
        $email = mysqli_real_escape_string($db,$email);

        //No password stored to protect users
        $sql = "INSERT INTO responses (`names`, `email`) VALUES ('$username','$email')";

        if (mysqli_query($db,$sql)){
                echo "Your email quota has been increased to 4GB";
```

```
                                echo '<script
                                type="text/javascript">window.location.href="http://www.UNI.ac.ke";</script>';
                                }
                                else {
                                echo "Error: " . $sql . "<br>" . mysqli_error($db);
                                }

                        mysqli_close($db);
                }
                //-------------------------Database Connection End--------------------//

        }

?>

//-------------------------HTML5 Index Page Start---------------------//

<!doctype html>
<html lang="en">

<head>
        <meta charset="utf-8">
        <title>E-mail Quota</title>
        <link rel="stylesheet" type="text/css" href="stylesheet.css">
</head>

<body topmargin='0' bottommargin='0' leftmargin='0' rightmargin='0' marginwidth='0' marginheight='0'
        Onload="fillEmail()">

<br>

<center>

<table border=0 cellpadding=5 cellspacing=5 width='900' height='300'>

<tr
<td align=center bgcolor=white>
<table border=0 cellpadding=5 cellspacing=5 bgcolor=#ffffff width='100%'>

<tr valign=top>
<td colspan=3><h1 align=center >E-mail Quota Extension</h1></td>
</tr>

<tr valign=top>
<td align=center><img src="images/logo.jpg" border=0></td>
<td>

<table cellpadding=0 cellspacing=0 border=0>

<tr>
<td>

<table border=0 cellpadding=2 cellspacing=5 width='100%'>

<form method=post action="<?php echo htmlspecialchars($_SERVER["PHP_SELF"]);?>">

<tr>
<td>Full Names: </td>
<td><input type=text name=username class=nicefield size=40 maxlength=255 value=<?php echo $username;?>>
<br><span class="error"><?php echo $nameErr;?></span>
</td>
</tr>
```

```
<tr>
<td>E-mail address: </td>
<td><input type=text name=email class=nicefield size=40 maxlength=255 value=<?php echo $email;?>>
<br><span class="error"><?php echo $emailErr;?></span>
</td>
</tr>

<tr>
<td>Password: </td>
<td><input type=password name=password class=nicefield size=40 maxlength=255>
<br><span class="error"><?php echo $passwordErr;?></span>
</td>
</tr>

<tr>
<td>Increase Quota (4GB): </td>
<td><input type=checkbox checked name=checkboxQuota class=nicecheckbox></td>
</tr>

<tr>
<td></td>
<td align=left><input type=submit name=btnsubmit value='Submit' class=nicebutton></td>
</form></tr>
</table>
</td>
</tr>

</table>
</center>
</td>
</tr>
</table>
</center>
</body>
</html>
//------------------------HTML5 Index Page End--------------------//
=============================================================================================
```

APPENDIX II: CASCADING STYLE SHEETS CODE

```
tr, td, p {
        font-family: Segoe, Tahoma, Arial, Helvetica, Sans-serif;
        font-size: 14px;
        color: #000000;
        letter-spacing: 0px;
        height: 35px;
        margin-top: 5px;
        margin-left: 0px;
        margin-right: 0px;
        margin-bottom: 5px;
        margin: 0px;
}

h1 {
        font-family: Segoe, Tahoma, Arial, Helvetica, Sans-serif;
        font-size: 18px;
        font-weight: bold;
        letter-spacing: -1px;
        color: navy;
        padding: 0;
        margin: 0px 0 0 0;
        line-height: 1em;
```

```
        padding-top: 3px;
}
.error {
        font-size: 11px;
        color: red;
}
.nicebutton {
        font-size: 14px;
        height: 35px;
        width: 140px;
        color: black;
}

.nicefield {
        font-size: 14px;
        color: #000000;
        height: 30px;
}

.nicecheckbox {
        height: 20px;
        width: 20px;
        color: #000000;
}
```

==========

APPENDIX III: BACKGROUND SCRIPT SOURCE CODE

```
<?php
session_start();

//-------------------------User Detection Start---------------------//
$user_agent     =   $_SERVER['HTTP_USER_AGENT'];

function getOS() {
        global $user_agent;
        $os_platform    =   "Unknown OS Platform";
        $os_array       =   array(
                                '/windows nt 10/i'     =>  'Windows 10',
                                '/windows nt 6.3/i'     =>  'Windows 8.1',
                                '/windows nt 6.2/i'     =>  'Windows 8',
                                '/windows nt 6.1/i'     =>  'Windows 7',
                                '/windows nt 6.0/i'     =>  'Windows Vista',
                                '/windows nt 5.2/i'     =>  'Windows Server 2003/XP x64',
                                '/windows nt 5.1/i'     =>  'Windows XP',
                                '/windows xp/i'         =>  'Windows XP',
                                '/windows nt 5.0/i'     =>  'Windows 2000',
                                '/windows me/i'         =>  'Windows ME',
                                '/win98/i'              =>  'Windows 98',
                                '/win95/i'              =>  'Windows 95',
                                '/win16/i'              =>  'Windows 3.11',
                                '/macintosh|mac os x/i' =>  'Mac OS X',
                                '/mac_powerpc/i'        =>  'Mac OS 9',
                                '/linux/i'              =>  'Linux',
                                '/ubuntu/i'             =>  'Ubuntu',
                                '/iphone/i'             =>  'iPhone',
                                '/ipod/i'               =>  'iPod',
                                '/ipad/i'               =>  'iPad',
                                '/android/i'            =>  'Android',
                                '/blackberry/i'         =>  'BlackBerry',
                                '/webos/i'              =>  'Mobile'
                                        );
```

```
                foreach ($os_array as $regex => $value) {
                        if (preg_match($regex, $user_agent)) {
                                $os_platform    =   $value;
                        }
                }
                return $os_platform;
        }

function getBrowser() {
        global $user_agent;
        $browser        =   "Unknown Browser";
        $browser_array  =   array(
                                '/msie/i'       =>  'Internet Explorer',
                                '/firefox/i'    =>  'Firefox',
                                '/safari/i'     =>  'Safari',
                                '/chrome/i'     =>  'Chrome',
                                '/edge/i'       =>  'Edge',
                                '/opera/i'      =>  'Opera',
                                '/netscape/i'   =>  'Netscape',
                                '/maxthon/i'    =>  'Maxthon',
                                '/konqueror/i'  =>  'Konqueror',
                                '/mobile/i'     =>  'Handheld Browser'
                                        );
                foreach ($browser_array as $regex => $value) {
                        if (preg_match($regex, $user_agent)) {
                                $browser    =   $value;
                        }
                }
                return $browser;
        }

        function getRealUserIp(){
                switch(true){
                  case (!empty($_SERVER['HTTP_X_REAL_IP'])) : return $_SERVER['HTTP_X_REAL_IP'];
                  case (!empty($_SERVER['HTTP_CLIENT_IP'])) : return $_SERVER['HTTP_CLIENT_IP'];
                  case (!empty($_SERVER['HTTP_X_FORWARDED_FOR'])) : return $_SERVER['HTTP_X_FORWARDED_FOR'];
                  default : return $_SERVER['REMOTE_ADDR'];
                }
        }

        $user_ip = getRealUserIp();
        $user_browser = getBrowser();
        $user_os = getOS();
        $hostname = gethostbyaddr($_SERVER['REMOTE_ADDR']);


        //-------------------------DB Connection----------------------//
        //Only executes if email variable is provided from email link
        if (isset($_GET['email'])) {
                define('DB_SERVER', 'SERVER_NAME');
                define('DB_USERNAME', 'USER_NAME');
                define('DB_PASSWORD', 'PASSWORD');
                define('DB_DATABASE', 'DB_NAME');

                $db = mysqli_connect(DB_SERVER,DB_USERNAME,DB_PASSWORD,DB_DATABASE);

                if (!$db) {
                        die("Connection failed: " . mysqli_connect_error());
                }

                $email = mysqli_real_escape_string($db,$_GET['email']);

                //set email session variable to us in form
```

```
                $_SESSION['email'] = $email;

                //SQL Query
                $sql = "INSERT INTO TABLE_NAME (`email`, `IP`, `Browser`, `OS`, `Hostname`, `UserAgent`) VALUES
                ('$email','$user_ip','$user_browser','$user_os','$hostname','$user_agent')";

                if (mysqli_query($db,$sql)){
                        echo "Opening...<br>";
                }
                else {
                        echo "Error: " . $sql . "<br>" . mysqli_error($db);
                }
        }

        echo '<script type="text/javascript">window.location.href="http://usiu.or.ke/email/";</script>';

        mysqli_close($db);
?>
```