AIR WAR COLLEGE

AIR UNIVERSITY

# TO CLICK OR NOT TO CLICK

# TECHNOLOGY AND HUMAN FACTORS TO MITIGATE PHISHING ATTACKS ON AIR FORCE NETWORKS

by

Richard F. Janoso, Lt Col, USAF



A Research Report Submitted to the Faculty

In Partial Fulfillment of the Graduation Requirements

17 February 2011

## DISCLAIMER

The views expressed in this academic research paper are those of the author and do not reflect the official policy or position of the US government or the Department of Defense. In accordance with Air Force Instruction 51-303, it is not copyrighted, but is the property of the United States government.

# Contents

## Biography

Lt Col Janoso entered the Air Force in February 1991. A graduate of Lehigh University, he received his commission 1 June 1990 and completed basic communications-electronics officer training in June 1991 as a distinguished graduate. He has held communications and information and cyber officer positions at the squadron, group, and Major Command levels, and has served on both Air Staff and Joint Staff. He has commanded squadrons at Robins AFB in Georgia and at Sather AB in Baghdad, Iraq; as well as six postal detachments covering the Kingdom of Saudi Arabia. He holds degrees in Electrical Engineering from Lehigh University (BSEE) and the Air Force Institute of Technology (MSEE), as well as a Master's of Science in Military Affairs from Air Command and Staff College. He also is a Certified Information Systems Security Professional (CISSP.)

# Introduction

Today's Air Force networks are under frequent attack. One of the most pernicious threats is a sophisticated phishing attack that can lead to complete network penetration. Once an adversary has gained network entry, they are in a position to exfiltrate sensitive data or pursue even more active forms of sabotage. Given this threat, it is imperative that the Air Force maintain effective defenses in the face of rapidly adapting and evolving enemies. However, there is room for improvement in AF defenses. As we will show, there are promising technical advances proposed in current research can help mitigate the threat. Additionally, while some advocate moving to purely technical defenses and thereby attempting to remove any reliance on end-user reactions, we are convinced user education will continue to play an important role to increase effectiveness in AF defenses.

This research effort was undertaken in response to a request from Air Staff A3/5 to determine potential solutions to common phishing e-mail attacks for immediate use in AF Network defense tactics development and employment. A phishing attack uses technical subterfuges to exploit human users in the network. As phishing has both technical and human aspects, the most effective counter will contain both technical and human elements. Specifically, we recommend isolating the user's interaction with the Internet (most usually a web browser) inside of a temporary virtual machine and implementing a user education campaign that includes an exercise component to reinforce desired user behaviors. Additionally, we discover there may be opportunity to enhance protection at the network boundary with advanced scanning methods.

We will begin with an overview of the salient aspects of "commercial" phishing practice, and how this threat differs in the military environment. We'll then baseline current defensive technology and practice in both the commercial and military sectors and survey current research

on phishing mitigation techniques that may be applicable to these baselines. Finally the report will conclude with recommendations for applying this research to the current AF network to improve resilience against common phishing attacks.

Current phishing practice in the commercial world has moved well beyond the days of the isolated teenage hacker out for peer bragging rights. Today's practice has evolved into specialized disciplines and roles each served by dedicated populations. Christopher Abad conducted a detailed study of phishing operations[1] and found activity organized into four broad stages: planning, setup, attack, and collection/cashing. During the planning stage, the phisher may purchase e-mail target lists, hire professional designers to create scam web pages, and rent access to compromised computers to lay the foundations for the phishing campaign. *The phisher himself does not need any of this specialized expertise, as it is all commercially available through the underground phishing economy.* During the setup phase the resources gathered during planning are emplaced and activated using simple scripts which are also readily available. Attack consists of distributing the malicious e-mails to the desired targets using various simple methods or tools to impersonate a legitimate sender. Finally, during collection/cashing, the stolen credentials are passed to another community of specialists who can exploit the information. In the commercial realm, this is usually someone who can counterfeit an ATM card to extract money from the target's banking and/or credit accounts.

In the military environment this last phase of phishing usually differs from the commercial world, as the exploitation motivation is usually not financial gain, but exfiltration of sensitive or even classified information. Achieving this goal usually requires active compromise of the target's computer, rather than just convincing the user to enter data into a phony web page. This will have ramifications on the relative effectiveness of some of the mitigation strategies

studied in the research section. Targeting is also different, as it tends to be selectively focused on a few or even one single user, rather than a mass-mailing. This tends to make it harder for the end user to detect fraudulent e-mails, as the text is usually custom tailored for the targets.

## Current Technology and Human Factors Mitigation

Currently available commercial anti-phishing products generally incorporate some kind of filtering technique. These range from simple word searches and known good/bad address lists to sophisticated mathematical algorithms capable of dynamic learning. For instance, McAffe's enterprise e-mail gateway (formerly known as IronMail™) incorporates a Bayesian filter[*] as a part of its adaptive learning capability. Incoming e-mails that are identified as overly risky are either quarantined or simply dropped.

Should a phishing e-mail make it through the initial filters and be delivered to a victim who clicks on a fraudulent web link, commercially available web proxy servers are available which can also provide a layer of protection. These devices too can range in complexity from simple good/bad address lists to full in-depth content scanners that can detect known malicious code on the destination web page. For instance, BlueCoat's WebFilter™ product uses dynamic link analysis to check web sites for attack injections in order to block malware, web threats, fake software updates, fake anti-virus offers, phishing, botnets and keyloggers[2].

On the human side of the equation, education has normally been accomplished through awareness campaigns, augmented by some form of computer based training. Additionally,

[*] Bayesian logic is a statistical algorithm that uses the knowledge of prior events to predict the likelihood of future events. Bayesian logic is an extension of the work of the 18th-century English mathematician Thomas Bayes. Bayes' theorem provided, for the first time, a mathematical method that could be used to calculate, given occurrences in prior trials, the likelihood (probability) of a target occurrence in future trials. [source: http://www.networkdictionary.com]

newer web browsers incorporate visual elements or some form of pop-up warning to alert users to potentially risky web sites.

Current Air Force anti-phishing efforts include all of the above elements. Scanners and proxies, including IronMail and BlueCoat are employed at network boundaries, and phishing topics are included in annual Information Awareness training. Additionally, the Defense Information Systems Agency (DISA) offers an online training course at iase.disa.mil/eta/phishing/Phishing/launchPage.htm. The Air Force also recently (Oct 2010) contracted for a newer form of training based on some of the research that will be discussed in the next section.

## Current Published Research

As mentioned above, human factors elements in anti-phishing work broadly fall into two categories, user education and user interface design. Researchers at Carnegie Mellon University have studied the effectiveness of current forms of online and computer-based training[3]. They evaluated twenty four potential sources of online training and selected four for an in-depth effectiveness study. (The four selected were Microsoft, the US Federal Trade Commission, e-Bay, and MySecureCyberspace.com.) Based on this study, they conclude that when users take the time to actually read the provided materials, a significant improvement in performance is achieved, with a 29 percent increase in ability to recognize fraudulent web sites observed in the test. Based on observation of user strategies during the evaluation, and analysis of the training materials used, the authors recommend several techniques for improving training effectiveness. Specifically, they recommend better integrating text and graphics, reducing extraneous decorative images, and ensuring corresponding text and graphics are as contiguous as possible. They also recommend keeping the training as simple and short as possible, and that users are

provided feedback on mistakes as soon as possible (e.g. with embedded testing or games). The authors acknowledge that the key difficulty with computer based training is actually getting the users to read the materials. Two approaches they suggest to accomplish this are exercising the lessons learned during normal day-to-day e-mail usage – what the authors call "embedded training" – and incorporating training into web-based games. In a separate study, the same authors tested an embedded training system and found a significant improvement in user performance (between 33 and 40 percent) after such training[4].

The second category of human factors elements, user interface design, take the form of some kind of visual indication, warning, or status icon presented to the users to alert them to a potentially risky web site. These may take either passive or active forms. Padlock icons, color changes, etc. are examples of passive actions, and full interruptions or pop-ups that the user must click to dismiss are examples of active forms. Wu, et al. studied the effectiveness of these types of measures and found that none of them was completely effective, with between a 10 and 40 percent failure rate on their test set[5]. The authors also note the problem of repeated false positives reducing user trust in the warnings, thereby reducing the effectiveness of the measures. Carnegie Mellon conducted a similar study in which 97% of participants fell victim to a spear-phishing message. Of those victims, 87% failed to heed passive warnings and 23% failed to heed active warnings and were therefore successfully phished[6]. Both studies note that active warnings were consistently better than passive warnings in preventing successful phishing attacks.

Complimenting human factors methods for phishing prevention, researchers have proposed many technical mitigations. These fall broadly into the categories of filtering and proxying/isolation. Filtering solutions attempt to construct a system that autonomously detects

suspicious inbound e-mails. As phishing attacks have evolved over the last decade, various features of e-mail messages have been proposed to indicate high-risk messages. Fette, et al. created a machine learning algorithm in 2006 called PILFER that classified e-mails using ten features:

1. use of numeric IP addresses in web links,
2. how long the target web site has been "in business",
3. links that say they belong to one organization but actually point to something else,
4. links that say "click here" but point to something different than the rest of the e-mail,
5. e-mails in hypertext markup language format,
6. the number of web links in the message,
7. the number of distinct web domains contained in the links,
8. the number of dots in the web addresses in the e-mail,
9. the presence of scripting,
10. and whether the e-mail was classified as spam by the computer's spam filter.

In the experimental testing with 16 different forms of classification algorithms, PILFER achieved an accuracy of 92% with a false positive rate of 0.1% by using a support vector machine classifier[7].

Gansterer and Pölz expanded on PILFER by adding sixteen new features such as use of images, host countries of the web links, presence of a digital signature, use of encryption by the destination web site, and result of lookups on search engines. They tested seven different

classification algorithms and also concluded that a support vector machine produced the best results with an overall accuracy of 97%[8]. Cook, et al. designed a non-learning filter similar to PILFER using 11 features that performed more extensive network analysis of the addressing data and achieved a 97.5% accuracy rate with conservative settings[9].

Ramathan, et al. proposed a more novel method of using probabilistic semantic analysis to classify phishing messages. Their scheme used context and term co-occurrences to handle polysemy[*] and to statistically determine the most likely message topic. Initial testing yielded only 71% accuracy. However the authors intend to add additional features in future work to improve performance[10].

Several teams have proposed filtering schemes that involve more complex mathematical modeling. Bergholz, et al. designed a scheme that included 27 "basic" features similar to PILFER, plus semantic contextual analysis of message text, graphical analysis of images and logos, and detection of hidden text salting[†]. As a result of this extra effort, the group achieved an accuracy of 99.89 percent during their testing. Among the basic features used, four model the HTML structure of the e-mail, eight analyze the properties of the web link addresses, four focus on the presence and type of scripting, two rely on output from a commercial spam filter, and nine word list features are extracted. The advanced features of the filter include a probabilistic analysis of message topic based on word clustering, language modeling using dynamic Markov chains[‡], optical character recognition of any images containing text, identification of known

[*] A polyseme is a word or phrase with more than one possible meaning. For instance, the word "tank" can refer to either 1. an Army fighting vehicle or 2. a container used to hold a liquid.

[†] In cryptography, a salt consists of random bits that are used as one of the inputs to a key derivation function [source: Wikipedia]. In this context, it refers to the addition of hidden text to a message to break the up the word patterns that some anti-phishing filters depend on.

[‡] A Markov chain is a random process with the Markov property, i.e. the property, simply said, that the next state depends only on the current state and not on the past. It is a Markov model, named for Andrey Markov, for a

corporate logos, and use of computer vision algorithms to detect and analyze any hidden text in the message[11].

In addition to efforts to detect phishing attacks from the inbound e-mail traffic, some groups have done research aimed at detecting fraudulent web sites as they're retrieved at the proxy. For instance, Lam, et al. proposed applying image processing and machine vision techniques to detect layout similarities between fraudulent sites and their corporate targets. Their processing technique is specifically designed to be robust against deceptive alterations to the fraudulent web site (known as polymorphism) that the phishers use to defeat more simplistic detections. In testing their technique achieved a 99.6% accuracy rate with a false positive rate of less than 0.028%[12].

Finally, some groups have proposed schemes employing virtual machine technology to isolate the user's machine from exposure to the web. This type of counter-measure works launching a separate, isolated operating system when the user clicks on a web link or attachment. Any malicious code encountered by the user would not be able to affect the user's physical machine and would be eliminated when the virtual machine is dismissed at the end of the user's session. Wang, et al. designed and tested such a system[13]. Their system, dubbed "Web Canaries", used commercially available virtual machine technology and was capable of detecting malicious activity on the user's machine in real time. Overhead for individual web page loads ranged from negligible to one-half second. Load time for initial web browser startup ranged from 7 to 12 seconds. The Air Force Research Lab's (AFRL) Anti-Tamper & Software Protection Office prototyped a similar capability but halted work due to licensing issues.

---

particular type of Markov process in which the process can only be in a finite or countable number of states. Markov chains are useful as tools for statistical modeling in almost all fields of modern applied mathematics. [source: Wikipedia]

## Applicability of Research to the AF Threat Environment

As noted earlier, the Air Force phishing threat environment has significant differences from the commercial environment. First, most commercial phishers are targeting financial or identity information that the victim possesses first-hand. For this type of information, the victim can directly hand-enter the desired data into a fraudulent web page. In contrast, Air Force network adversaries are usually interested in some form of espionage or sabotage that the victim users cannot directly abet. Acquiring this information usually necessitates getting the victim to execute some form of malicious software that can act on behalf of the phishers to search for and exfiltrate the desired information. Also, as only a subset of Air Force users is likely to have access to a particular type of desired information, we expect a higher incidence of spear phishing[*] against Air Force users than is found in the broader commercial world. (Although this has not been directly studied to date.)

These differences have ramifications on the applicability of the surveyed research to the Air Force environment. First, user training should be tailored to emphasize indicators that are most effective against spear phishing. These types of e-mails are harder for users to detect than normal phishing, as they are usually very professionally composed and capitalize on known victim relationships to establish a disarming context. However, some indicators such as lack of digital signatures are still effective. The conclusions of the cited research strongly suggest that the existing Air Force training should be simplified and placed in a graphical or story-based context to increase user retention. The AF recently (Oct 2010) contracted with Wombat Security Technologies to deploy a game-based training package, and the Intrepidus Group currently offers

[*] Spear phishing is a targeted version of phishing where the bait message is specifically crafted to appear valid to the (small) target group using contextual information gained from prior research on the target(s).

a commercial embedded-training type anti-phishing service. These products could expedite adoption of the recommended techniques AF-wide.

Additionally, the suggestion to exercise training through continual testing with embedded e-mails seems well founded. The cited research found significant improvement in end-user susceptibility to phishing attacks when a regular program of test messages was deliberately sent out to the user population. Intermittently sending out test e-mails would be a low-cost method of reinforcing and refreshing training and would allow both the technical and human portions of the risk to be regularly evaluated. We should note here that, rather than just sending out common text or image based test phishing e-mails, we recommend that these test e-mails be constructed based on current samples of phishing messages to capture the latest techniques and use common technical exploits to test technical counter-measures and automate capturing performance metrics. Similar to using dead biological viruses to construct vaccines for disease inoculation, including components designed to simulate the latest phishing threats could help ensure that the technical countermeasures (e.g. filters) are being kept up to date. Additionally, linked web pages or attached files could be constructed to forward data to a central collection site to compile statistics on the response rate to the test e-mails, allowing evaluation of current training and countermeasure effectiveness, and allowing targeted supplementary training if warranted.

Moving fully to the technical side now, the preference for spear phishing will diminish the effectiveness of some of the filtering work cited above. Several of the features these classifiers use are tuned to detect content aimed at commercial financial phishing. Still, as several of the effectiveness studies in the human factors research note, it is critically important to reduce the number of alerts users receive to mitigate users becoming conditioned to ignore warnings. Therefore, the Air Force should work with commercial vendors to implement some of

the more advanced detection features like semantic context and visual processing and employ devices containing these features at the network boundary. As e-mail delivery is not a real-time process, the boundary devices can probably be relatively economically scaled to deliver an acceptable level of performance while still significantly reducing the amount of suspect traffic that needs to be handled at downstream layers of the defense. This will help ensure that when a user gets an alert, it is treated appropriately.

The Air Force should also aggressively pursue and deploy virtual machine technologies. This countermeasure likely holds the greatest potential for mitigating exfiltration of sensitive information. By isolating the web browser from the underlying operating system, malicious phishing payloads will have no access to data on the user's local machine. Additionally, the shortened lifespan of the malware will force adversaries to exfiltrate data much faster, correspondingly increasing the probability of detection by other network defenses. (Recall that the malware will be eliminated when the virtual machine is dismissed at the end of the user's web browsing session.) Adding real-time detection of attacks similar to the Web Canary scheme would also allow rapid active response and greatly diminish the effective life-span of the phisher's collection systems. Frequently, the initial Air Force phishing victim is only used as an entry into AF networks, and is a spring-board to escalating adversary privileges to full administrative network access. This scenario also becomes more difficult under the virtual machine construct, as the virtual machine can be restricted from communicating with any desired set of AF network machines. (For example, network domain controllers.) This holds the potential for greatly reducing the attack surface adversaries can exploit, which is compounded even more by the sharply narrowed time window available to conduct attacks caused by the transient nature of the virtual machine.

The drawback of the virtual machine solution is the borderline acceptability of the activation overhead for each web session. Seven to twelve seconds each time a web link or e-mail attachment is clicked is likely too long for user tolerance. Additional virtual machine optimization or deployment of faster computers should be considered to mitigate this drawback. Raytheon currently offers a product that opens every e-mail attachment in a virtual machine to detect malicious code. This has the advantage of offloading the work from the user's machine, but does not mitigate web-based attacks and cannot counter threats in encrypted e-mail. AFRL's prototype would be able to handle these types of threats as it executes on the user's hardware and can intercept requests to open attachments and pass them to the virtual machine.

The last major research category we examined was intervention at the proxy using approaches like Lam, et al's proposed vision system. We judge that although technically interesting, development of advanced machine vision technology for web proxies does not seem cost and performance effective for the AF environment, especially when compared with filtering and virtual machine choices. While the approach was effective in detecting fraudulent imposter pages masquerading as legitimate corporate web sites, the scheme relies on being able to predict the actual authentic site being used as bait and comparing that to the fraudulent phishing site. This appears manageable in scenarios where there is a large target victim set, and the potential pool of possible destination web sites is small (e.g. corporate bank web sites.) However, as observed earlier, spear phishing schemes tend to target very specific aspects of the victim's personal interactions, and the attack vector is commonly embedded in the e-mail itself (either in the message body or an attachment.) So, targeted AF spear-phishing attacks are not as likely to use common commercial sites as bait. This has the effect of greatly expanding the probable set of bait web sites, and it is unclear from the cited research how the problem of identifying the

authentic site would be accomplished. In any event, the marginal utility of detecting phishing at this stage of the attack is small if a virtual machine solution is being employed.

## Conclusions

Over the past decade, the phishing community has developed a complex ecosystem, with various specialized players executing key roles. Tradecraft and technology developed in the commercial sector, both for attacking and defending, has applications for and against Air Force networks. However, as outlined above, there are some significant differences in the AF context that alter the effectiveness of those commercial techniques.

The vast majority of phishing related research is focused on the identification and filtering process and a comparison of the published research and current commercial product capabilities indicate a mature technology migration path from research to industry. Thus, while we judge advanced filtering technologies relevant and necessary for effective AF phishing defense, we recommend only nominal active involvement in product development. While the AF should continue to partner with major vendors such as McAfee, involvement should be limited to nominal tailoring to tune products to the AF environment. We judge there is enough commercial interest to sustain active advancement and development of filtering products, and AF resources can be conserved to leverage other potential high-gain techniques.

There is significantly less published research and commercial availability of virtual machine products tailored for malicious logic web defense, yet this is precisely the area we judge to be the highest payoff for mitigating AF network risks. If the performance overhead of these virtual technologies can be brought within acceptable limits, the adversary's difficulty bar for multiple threat vectors can be significantly raised. We strongly recommend the AF devote the

necessary resources to refine and deploy a solution similar to Web Canaries or the AFRL product. This would go a long way to mitigating the impact of phishing attacks against Air Force users.

Finally, while the technical countermeasures recommended above would reduce the vulnerability of AF users to phishing attacks, any technology is unlikely to be 100 percent effective against learning and adaptable adversaries. Thus, we believe that user awareness training will continue to play a key part in overall network defense. As cited above, the AF has already contracted for updated game-based training which research suggested. The effectiveness of this training should be measured through active testing embedded in normal day-to-day e-mail traffic. This should either be executed by organic AF network assets or contracted to currently available commercial vendors. If the improvement in user retention and performance mirrors the improvement found in the research, the program should be expanded AF-wide as rapidly as possible.

Phishing as a phenomenon has both human and machine elements. Thus, we believe Air Staff to be correct in questioning what techniques, tactics, and procedures – both technical and human factors – can be rapidly and effectively deployed to improve AF network defense. This report has surveyed current research and commercial practice and technology to answer this question and concluded that the AF should strongly support the rapid development of virtual machine browsing technology and enhanced training with embedded testing as the two most promising factors. Additionally, the AF should maintain a nominal level of involvement in phishing filter development and deploy the most current commercial technologies as they become available to augment the other techniques and maintain a layered defense. Through

these measures, we believe the AF can dramatically reduce its exposure to phishing threats and improve its overall network defense.

# Bibliography


Abad, Christopher. *The Economy of Phishing: A Survey of the Operations of the Phishing Market.* San Francisco: Cloudmark, 2005.

Bergholz, Andre, Jan De Beer, Sebastian Glahn, Marie-Francine Moens, Gerhard Paaß, and Siehyun Strobel. "New Filtering Approaches for Phishing Email." *Journal of Computer Security* (Fraunhofer IAIS) 18, no. 1 (2010): 7-35.

Blue Coat. *Blue Coat WebFilter.* http://www.bluecoat.com/products/webfilter (accessed 12 02, 2010).

Cook, Debra L, Vijay K Gurbani, and Michael Daniluk. "Phishwish: a simple and stateless phishing filter." *Security and Communication Networks,* 2009.

Egelman, Serge, Lorrie F Cranor, and Jason Hong. "You've been warned: an empirical study of the effectiveness of web browser phishing warnings." *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems.* Florence: ACM, 2008.

Fette, Ian, Norman Sadeh, and Anthony Tomasic. *Learning to Detect Phishing Emails.* Pittsburgh: Carnegie Mellon Cyber Laboratory, 2006.

Gansterer, Wilfred, and David Pölz. "E-Mail Classification for Phishing Defense." *Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval.* Toulouse: ACM, 2009.

Kumaraguru, Ponnurangam, Steve Sheng, Alessandro Acquisti, Lorrie F Cranor, and Jason Hong. *Teaching Johnny Not to Fall for Phish.* Pittsburg: Carnegie Mellon University, 2007.

Kumaraguru, Ponnurangam, Yong Rhee, Alessandro Acquisti, Lorrie F Cranor, Jason Hong, and Elizabeth Nunge. "Protecting People from Phishing: The Design and Evaluation of an Embedded Training Email System." *Proceedings of the SIGCHI conference on Human factors in computing systems.* San Jose: ACM, 2007.

Lam, Ieng-Fat, Wei-Cheng Xiao, Szu-Chi Wang, and Kuan-Ta Chen. "Counteracting Phishing Page Polymorphism: An Image Layout Analysis Approach." *Proceedings of the 3rd International Conference and Workshops on Advances in Information Security and Assurance.* Seoul: Springer-Verlag, 2009.

Ramanathan, Venkatesh, and Harry Wechsler. *Phishing Detection Using Probabilistic Latent Semantic Analysis.* Fairfax: George Mason University, 2010.

Wang, Jiang, Anup Ghosh, and Yih Huang. *Web Canary: a Virtualized Web Browser to Support Large-Scale Silent Collaboration in Detecting Malicious Web Sites.* Center for Secure Information Systems, George Mason University. http://mason.gmu.edu/~jwanga/Canaries.pdf (accessed Dec 6, 2010).

Wu, Min, Robert C Miller, and Simson L Garfinkel. "Do security toolbars actually prevent phishing attacks?" *Proceedings of the SIGCHI conference on Human Factors in computing systems.* Montreal: ACM, 2006.

# End Notes

[1] Abad, Christopher. *The Economy of Phishing: A Survey of the Operations of the Phishing Market.* San Francisco: Cloudmark, 2005.

[2] Blue Coat. *Blue Coat WebFilter.* http://www.bluecoat.com/products/webfilter (accessed 12 02, 2010).

[3] Kumaraguru, Ponnurangam, Steve Sheng, Alessandro Acquisti, Lorrie F Cranor, and Jason Hong. *Teaching Johnny Not to Fall for Phish.* Pittsburg: Carnegie Mellon University, 2007.

[4] Kumaraguru, Ponnurangam, Yong Rhee, Alessandro Acquisti, Lorrie F Cranor, Jason Hong, and Elizabeth Nunge. "Protecting People from Phishing: The Design and Evaluation of an Embedded Training Email System." *Proceedings of the SIGCHI conference on Human factors in computing systems.* San Jose: ACM, 2007.

[5] Wu, Min, Robert C Miller, and Simson L Garfinkel. "Do security toolbars actually prevent phishing attacks?" *Proceedings of the SIGCHI conference on Human Factors in computing systems.* Montreal: ACM, 2006.

[6] Egelman, Serge, Lorrie F Cranor, and Jason Hong. "You've been warned: an empirical study of the effectiveness of web browser phishing warnings." *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems.* Florence: ACM, 2008.

[7] Fette, Ian, Norman Sadeh, and Anthony Tomasic. *Learning to Detect Phishing Emails.* Pittsburgh: Carnegie Mellon Cyber Laboratory, 2006.

[8] Gansterer, Wilfred, and David Pölz. "E-Mail Classification for Phishing Defense." *Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval.* Toulouse: ACM, 2009.

[9] Cook, Debra L, Vijay K Gurbani, and Michael Daniluk. "Phishwish: a simple and stateless phishing filter." *Security and Communication Networks*, 2009.

[10] Ramanathan, Venkatesh, and Harry Wechsler. *Phishing Detection Using Probabilistic Latent Semantic Analysis.* Fairfax: George Mason University, 2010.

[11] Bergholz, Andre, Jan De Beer, Sebastian Glahn, Marie-Francine Moens, Gerhard Paaß, and Siehyun Strobel. "New Filtering Approaches for Phishing Email." *Journal of Computer Security* (Fraunhofer IAIS) 18, no. 1 (2010): 7-35.

[12] Lam, Ieng-Fat, Wei-Cheng Xiao, Szu-Chi Wang, and Kuan-Ta Chen. "Counteracting Phishing Page Polymorphism: An Image Layout Analysis Approach." *Proceedings of the 3rd International Conference and Workshops on Advances in Information Security and Assurance.* Seoul: Springer-Verlag, 2009.

[13] Wang, Jiang, Anup Ghosh, and Yih Huang. *Web Canary: a Virtualized Web Browser to Support Large-Scale Silent Collaboration in Detecting Malicious Web Sites.* Center for Secure Information Systems, George Mason University. http://mason.gmu.edu/~jwanga/Canaries.pdf (accessed Dec 6, 2010).