Phishing in 2012 and Beyond:

The past, present, and future of the Internet's most lucrative scam

Aaron K. Nelson

COMM 610*301

Introduction

The Internet is quite a bit like cinema's depiction of the Wild West. With a great deal of optimism, society has followed the path of a few brave pioneers and set out to cultivate the digital frontier. Like the Old West, the Web presents a vast potential that we are only beginning to realize. If it is to flourish, this new environment requires personal responsibility and good citizenship, since law enforcement can hardly keep up with the exponential pace of online growth—and thus, the online frontier presents new opportunities for bandits and criminals.

The threats are pervasive. A typical user could be sharing their most personal information with companies who care very little about protecting it, divulge private and intimate material to parties who might exploit it, have his or her computer enslaved into doing the nefarious bidding of a new master, or become a new hub for snatching up additional victims.

But just as a townsperson in the Wild West could not completely avoid gunfights by abstaining from the saloon—Hollywood says most gunfights seemed to happen in the streets—the diligence required to avoid all manner of possible online attacks goes deeper than just steering clear of the dark and scary corners of the Internet. One of the most fruitful methods at the disposal of the malicious is to create websites, e-mails, and social media posts that are ostensibly from legitimate, credible entities, such as banks or retailers, in order to dupe users into willingly revealing information or surrendering their money. This process constitutes *phishing*.

Phishing is now a constant threat. Every opened e-mail and every clicked link could be leading a user to a well-crafted imitation whose only purpose is siphoning off sensitive data. Roughly 1,000 new attacks are created every month, and tens of millions of Americans have fallen prey to them so far (Bindra, 2010, p. 366).

With phishing so rampant, we as a society need to better understand the underlying causes as to why it is still so successful. If we can do that, we can improve our efforts from every angle: from public awareness and education, to reporting of phishing attacks, to law enforcement and punishment for phishers. To defeat an enemy, one must first understand the enemy, and so it is germane to research the question:

How are scammers using words, styles, elements, and other features to lend credibility to their phony websites?

If we can comprehensively answer that question, we are off to a strong start. Just as a malicious user must first understand a system in order to exploit its vulnerabilities and attack it, we need a deep understanding of the methods being employed (and the reasons those methods are working) if we are to educate the public about how better to avoid, report, and punish those responsible. The end goal—a safer Internet—is a noble pursuit.

Method

To find out how scammers are doing their iniquitous work, there is little need to go seeking and analyzing spam and fake websites firsthand. A great deal of research has already been conducted on the subject, pulling phishing attacks from the wild for digestion and deconstruction.

This paper will take a deeper look at the basic phishing process through the analysis of academic research, primarily conducted by professionals in the information technology, information security, and media industries. It will cover how phishing works and why it works, as well as include some more advanced and sophisticated methods and tricks along the way. Then, it is necessary to cover the remedies that have been used to try and thwart such attacks, and how well they have performed. It will conclude with an overview of some of the emerging new ideas that could present real solutions, and suggest direction for further research.

Results

A Brief History

At their heart, most Internet scams did not spring up out of thin air with the advent of the World Wide Web. Rather, they can trace their heritage back to scams that have existed for hundreds of years. Both online and offline grifts are successful because they exploit flaws in us that have always been a part of mankind.

Frank Stajano and Paul Wilson (2011) outlined these vulnerabilities and boiled them down to seven major weaknesses. Computer users, and people in general, are: easily distracted (they tend to focus on what they are interested in, ignoring security features aiming to help them and red flags that would otherwise alert them to danger); they are conditioned to obey authority, and will more readily comply with unreasonable demands when they seem to be coming from someone trustworthy; they are herd-minded, and will lower their defenses when those around them are doing the same; they are at times dishonest, willing to bend morals for a quick and easy gain (and afterwards are even more compliant, due to the possibility of blackmail); they are also fundamentally nice and generous, qualities which one might take advantage of; they have basic needs and desires which can be exploited; and, when under the stress of a limited time or a deadline, make decisions more hastily. Stajano and Wilson posited that every one of these qualities, which still sucker tourists into playing Three-card Monte and con jewelry store clerks into giving up merchandise to fake cops, are the same reasons users fall for online scams.

The Basic Phish

The basic flow of a typical phishing scam goes something like this: a malicious user creates a fake web page purporting to be that of a legitimate business—most frequently a bank, online retailer, Internet Service Provider, or government—using the design of the actual website as source material (Lynch, 2005, p. 267). They create forms to capture information—from personal details which can be used to steal an identity to financial details to be used for a quick cash grab. They host it online, often choosing domain names (the website's address) that plays off of the original. Users who are negligent, distracted, or just uneducated may not notice that, when browsing to or clicking links saying they lead to "example.com", they could actually be common phishing knock-offs like:

- a new domain name close to the spoofed original, perhaps one letter off (exammple.com)
- a new domain name that plays on keyboard layout and common typos (exzample.com)
- a different domain name with the spoofed address included (example.com-malicious-site.com)
- a different domain name with the spoofed address tacked on at the end ("malicious-site.com?site=example.com)

But rather than wait for traffic to stumble across this fake site, phishers must actively seek users who are especially vulnerable. They compose an e-mail posing as the entity, demanding some call to action. The techniques here sometimes border on artistic. Traditionally, they may talk about some problem with your account, a need to verify information, perhaps even that you are at risk of losing your account or your money. More advanced phishers—Hal Berghel (2006) calls them "phish mongers"—compose more advanced e-mails. Since many e-mail providers scan the text of a message for clues that it could be a scam, they might create the entire body of the message as an image, and include a modicum of benign-looking nonsense text, made invisible by writing in white on white, to do an end-run around the filter (p. 24).

If the phisher spams this message out to millions of people, it will surely reach thousands of users who do business with that entity, and it is likely to catch a few suckers. Once on the fake site, any information provided is captured and stored by the phisher. Another advanced trick is to allay any suspicions the user might have—after all, what good is a credit card or bank account if the owner gets wise and closes it down? In order to avoid detection, the phisher is likely to redirect to the entity's actual website, perhaps even passing along any information their victim has supplied (Berghel, 2006, p. 23).

E-mail was once the only place phishers could anonymously reach out to millions of people, but as social media pervades every aspect of our lives, it also provides a growing platform for phishers to cast their lines. Bamnote, Patil and Shejole (2010) conducted a study to find how easily they could direct traffic to a given website using fake accounts set up on Facebook. By simply using a photo of an attractive girl and adding complete strangers, they were able to accumulate hundreds of "friends." In a twist, they tried adding the friends of their friends, who would then be notified that this was not a total stranger, but rather a mutual friend—and got a 60% acceptance rate. Most importantly to the study of phishing, however, when they posted a link using these completely bogus accounts—a link whose real address was hidden with a free URL-shortening service—28% of their new "friends" clicked it (p. 153). If a few researchers are able to see results like that, it is reasonable to assume a dedicated phisher can adapt and fine-tune this method, and find good results without ever resorting to spamming e-mail.

Basic Countermeasures

From the beginning, steps were taken to combat the emergence of phishing schemes. One of the earliest ideas, and one still in use today, is the TRUSTe seal, designed to give users a quick overview of how a website collects your information and what measures it takes to keep it secure. With one click, you could be taken to the TRUSTe website where, in addition to learning about how that page handles your privacy, you could be assured that you were referred there by the legitimate page (Benassi, 1999, p. 59). Since its foremost purpose was keeping online companies honest about their information-gathering techniques, and preventing spoofing was never a priority, one has only to look at the continued rampancy of phishing to see that this was never effective. Users may have ignored the practice, taking it for granted that the seal alone was enough to certify a site without clicking it to verify.

There are also numerous avenues to report phishing. Google, eBay, the IRS, large banks, the federal government—essentially every entity at serious risk of being spoofed in a phishing attempt, as well as most mail clients and several third-party services—all offer a means for reporting phishing attempts. Their efficacy is a mixed bag: it may be a simple process to have the phishing site's registrar or host deactivate the account, but the process is a bit like fighting an ant infestation with only a hammer. These agencies cannot fight a thousand new threats every month based only on the few reports made by concerned Samaritans.

It did not take long for phishers to find that the anti-fraud departments of businesses make the perfect bait to phish with. What better way to lower a victim's guard than to say "you have been phished, click here to dispute this charge…"? Since spoofed websites would be perfect clones of the originals, Diniev (2006) cites a massive April 2004 phishing campaign, whose faked Paypal pages included links to "'Report a Spoof' and 'Avoid Fake Web sites'" (p. 80).

The biggest weapon in the arsenal in the fight against phishing has always been education. Lynch (2005) cited the Federal Trade Commission, who found in a survey that victims of identity theft felt that they would have been better equipped to deal with the crime if they had "better awareness" (p. 276). She elaborates, however, that there is a paradox in asking consumers to self-educate:

> Another somewhat obvious problem is that the term "phishing" is so obscure that many consumers would not associate it with a suspect e-mail… Therefore, consumers face a Catch-

22—they could prevent a successful phishing attack if they knew what to look for, but without knowing what to look for (i.e., that the scam is called "phishing"), consumers may be hard-pressed to educate themselves about the attacks.

The next logical step, then, is to take the initiative in educating the masses via public service announcements, but PSA campaigns also have their limitations. First, spotting phishing attempts (especially the well-crafted ones) can be difficult even for very savvy users. Getting good at spotting these scams requires in-depth, technical education about how web browsing really works—a process that cannot be covered in a 30-second radio spot.

Advanced countermeasures

Today, a great many researchers are hard at work on exciting new ways to use emerging technologies in the fight to identify online scams before a victim can fall prey.

Every day, as a part of their design, the algorithms that scan e-mail and the databases that track e-mail servers notorious for spam get smarter and stronger. Self-teaching machines may be our best hope. Rather than compile a list of known fake sites, a process which is inherently one step behind, there are designs in the works that aim to use statistical learning theory (in essence, machines that learn as they go) to spot fake sites by analyzing them from every angle—the colors of every pixel of every image, the structure of the page's source code, the use of language and grammar, the URL itself and its construction—and determine whether it is real or fake before being displayed to a potential victim (Abbasi et al., 2005).

One of the most common targets of a phishing attack is simply login information, with which an attacker can compromise the rest of an account and access any information stored there. Mannan and Oorshot (2011) advocate a whole new approach to the age-old login system, a process that is slowly being adopted by big sites like Google and Facebook. Rather than simply supply a username and password, which can be compromised, you link your account to a real-world mobile device, which you physically own and carry. Upon trying to log in with a name and password, the site then sends your device a short number, which you then supply, thus authenticating that it is really you logging in. The number is valid for only one use, and expires quickly. Setups exist for this number to be texted to a mobile phone, or to be transmitted to a simple, inexpensive new device with the sole purpose of receiving and displaying these numbers.

Conclusion

From the body of knowledge provided by the research, a few basics are clear. Phishing is a serious, ubiquitous crime, with costs to consumers numbering in the billions of dollars and costs to businesses in the tens of billions (Lynch, 2005, p. 261).

By utilizing the body of research to analyze the phishing process, and by further dissecting it, solutions begin to emerge. Because the perpetrators are so good at adapting and overcoming roadblocks thrown in their path, combating the problem must require a multifaceted approach. The first line of defense will always be public awareness and education. If a user is wise to these tricks, no harm is done.

How phishers operate—the words, styles, elements, and other features they use to lend credibility to their phony websites—goes much deeper than parroting the original site's look and feel. They take psychological advantage of a person's trust in authority, distracting and misdirecting like a stage magician, and employing a panoply of tricks to get their message seen and believed.

There is an inherent difficulty in analyzing phishing techniques and conceiving solutions, because it is constantly evolving. Researching and understanding phishing is a process, one which neither this nor any other paper can exhaustively cover.

Phishing has always been an exploit of vulnerabilities in the way the Internet works—how links are constructed, how e-mail is sent, and how we transmit our information. Legitimate businesses have always been playing catch-up, reacting to the new moves of phishers. To truly make a change in the problem, we must fundamentally rethink our methods and perhaps even fight fire with fire; some of the most exciting and promising new ways of coping with the risks created by our technologies are newer, smarter technologies. As phishers remain on the cutting edge of technological prowess, exploiting every possible trick to make their craft appear legitimate, the next step in fighting it must involve putting some of that technological advantage back in the hands of the user, by deploying not only smarter machines, but machines that can learn. There is already something of an arms race between the ability to detect malicious spam and the petty scammers' ability to juke around these detections. In the future, the emergence of artificial intelligence may tip the scales, as the guardians of our inboxes not only learn what new tricks are in use and how to see them, but begin to understand the art of the scam.

References

Abbasi, A., Zhang, Z., Zimbra, D., Chen, H., & Nunamaker, J. F. (2010). Detecting fake websites: The contribution of statistical learning theory. *Management Information Systems Quarterly, 34*(3), 435-461. Retrieved from http://www.misq.org/contents-34-3/

Bamnote, G., Patil, G., & Shejole, A. (2010). Social networking: Another breach in the wall. *American Institute of Physics Conference Proceedings, 1324*(1), 151-153. doi:10.1063/1.3526180

Benassi, P. (1999). TRUSTe: An online privacy seal program. *Communications of the Association for Computing Machinery*, *42*(2), 56-59. Retrieved from http://cacm.acm.org/magazines/1999/2/7971-truste/

Berghel, H. (2006). Phishing mongers and posers. *Communications of the Association for Computing Machinery*, *49*(4), 21-25. Retrieved from http://cacm.acm.org/magazines/2006/4/5936-phishing-mongers-and-posers/

Bindra, G.S. (2010). Efficacy of anti-phishing measures and strategies: A research analysis. World Academy of Science, Engineering & Technology, 69366-372. Retrieved from https://www.waset.org/journals/waset/v45/v45-66.pdf

Dinev, T. (2006). Why spoofing is serious Internet fraud. *Communications of the Association for Computing Machinery*, *49* (10), 77-82. Retrieved from http://cacm.acm.org/magazines/2006/10/5805-why-spoofing-is-serious-internet-fraud/

Lynch, J. (2005). Identity theft in cyberspace: Crime control methods and their effectiveness in combating phishing attacks. *Berkeley Technology Law Journal, 20*(1), 259-300. Retrieved from http://btlj.org/category/journal/

Mannan, M., & van Oorschot, P. C. (2011). Leveraging personal devices for stronger password authentication from untrusted computers. *Journal of Computer Security, 19*(4), 703-750. doi:10.3233/JCS-2010-0412

Stajano, F., & Wilson, P. (2011). Understanding scam victims: Seven principles for systems security. *Communications of the Association for Computing Machinery, 54*(3), 70-75. doi:10.1145/1897852.1897872