Web Security Requirements: A Phishing Perspective

Ian Fette
School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
Email: icf@cs.cmu.edu

Norman Sadeh
School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
Email: sadeh@cs.cmu.edu

Lorrie Cranor
School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
Email: lorrie@cs.cmu.edu

***Abstract*— We are currently focusing on web security problems caused by phishing, and similar semantic attacks against users. Our current investigations are leading towards heuristic, collaborative, and semantic approaches towards thwarting such attacks. Additionally, we are considering new approaches to authentication that minimize the room for user error in the presence of semantic attacks. We feel that there is significant room for progress in both of these areas, and that further testing to validate any potential solution to web security problems must take semantic attacks into account in the context of real user behavior.**

I. Introduction

Phishing is a growing problem[1] that affects an increasing number of users and companies providing online services. At its most fundamental level, phishing is a subset of a larger class of semantic attacks against the user, which are seen as defining the current wave of network attacks[2]. These attacks are increasingly perpetrated by targeting the user's environment and interfaces. A lack of usable mutual authentication opens consumers up to both classic man in the middle attacks and semantic attacks, showcasing the need for a solution to help users authenticate service providers in a usable and secure manner.

We believe that there is potential for immediate return on technologies designed to detect spoofed webpages and emails. To that end, we are currently investigating heuristic approaches and collaborative approaches designed with the goal of determining the authenticity of an email or webpage. We hope that this will offer a sufficiently effective approach to thwart the majority of phishing attacks. We are also considering a few long-term approaches, including semantic reasoning and analysis of attacks, as well as examining the fundamental faults in authentication mechanisms that make these attacks possible. Some solutions we are currently considering include leveraging the pre-existing out of band communications that take place between customers and institutions.

II. Detecting spoofed content

A. Heuristic approaches

There are a number of heuristic approaches that we believe may be effective in detecting phishing attacks, starting at the email level. These approaches are based on an understanding of common traits found in phishing emails, as described in [3]. Besides the traditional spam filtering approaches, such as Bayesian filters, we believe that filters acting on certain key characteristics may be able to filter out many phishing attacks at the email level. Such characteristics include emails with links to IP addresses, emails with links to newly registered domains, and emails that appear commercial in nature but originate from either residential or foreign IP addresses.

Heuristics are already in use in anti-phishing toolbars, such as Spoofguard[4] and the Netcraft Toolbar[5], and we are currently conducting studies to evaluate the effectiveness of these various approaches. Our preliminary studies indicate that the accuracy of currently available toolbars varies quite a bit from product to product. Unfortunately, applying heuristics at the email level must rely on different techniques than applying heuristics at the browser level.

At the most basic level, different information is available in the email (such as header information), but on a higher level, the cost of heuristics at an email level must be weighed differently. Any filtering done by an ISP, or similar email gateway, must have a low marginal processing cost due to the high volume of mail to be screened. Filtering done in the user's web browser can have a relatively higher marginal cost, since there is far less volume, and less sensitivity by humans to small (sub-second) delays. These small delays can quickly add up and halt an email server, however, if 100 emails per second arrive with each email taking more than 1/100s to process.

B. Collaborative approaches

Given that there are on the order of thousands of phishing attacks per month[6] going to hundreds of millions of users, it seems reasonable to hope that a collaborative approach could also assist in detecting a large number of phishing attacks. A typical phishing attack seems to be sent out multiple times to many different people, and if the first recipient recognizes the email as a phishing email, it should be possible for that person to effectively inform the community, acting in effect as a vaccine against further emails of a similar nature.

For this to be feasible, a number of new developments are required. First, there needs to be a framework for a scale-free network to communicate such "vaccination" information among peers. Second, there must be a way of accurately developing such information such that a maximum number of similar attacks are matched, but legitimate emails are not matched. Third, there needs to be a way to evaluate and disseminate information about the quality of such reports, to prevent abuse.

C. Semantic approaches

Semantic attacks have existed since long before the Internet. As phishing attacks become increasingly sophisticated, it may become necessary to fight them on a semantic level. If, for instance, my mail client knows that I do not have an account with Bank XYZ, then an email telling me that I need to update my account information at Bank XYZ should be disregarded. It might also be possible for my client to recognize that I do indeed have an account with Bank XYZ, but that in the past I have communicated with Bank XYZ at xyz.com and that an email appearing to be from Bank XYZ but asking me to log in to ZYX.com should be treated with suspicion. The basic insight behind semantic approaches is to build a deeper understanding of a user's activities and the entities (s)he interacts with and to see whether such an understanding can help detect potential attacks. Phishing attacks reflect the increasing amount of personal data available and transmitted online. Semantic approaches to combating phishing attacks aim at leveraging this very data. Such approaches could be embedded in the form of intelligent software agents that combine features of digital wallets or federated identity management systems with sophisticated knowledge representation, data mining and reasoning functionality (e.g. see [7] for an early example).

III. Authentication Considerations

A. Leveraging Out-Of-Band Communications

Many security technologies exist today that, in theory, would solve the problem of strong, mutual authentication. In an ideal world, we could have one huge PKI infrastructure with well-defined trust models, and I would know that I am communicating with my bank and vice versa. In reality, PKI implementation has turned out to be a daunting task, and many question whether a global implementation is achievable.

We are currently considering the ramifications of taking a step back from the ideal global infrastructure, and looking at what is possible without the existence of a clear trust hierarchy. In short, what you are left with is a set of private and public keys, and the same out-of-band identification mechanisms present in today's society. We are considering ways to leverage the existing out-of-band authentication to support a setting with pre-shared keys, including technologies to facilitate such an environment.

In short, we propose considering PKI as simply a private-public keypair, leveraging the communications security therefrom, and leaving the association of a key to an individual in the realm of current identification practices. This has the same risk factors as current establishment practices for a new relationship (a consumer with a new bank, a bank with a new customer, for example) but may provide much stronger security for subsequent interactions than is currently feasible.

B. Usability and User Testing

Unfortunately, usability is frequently given insufficient consideration with respect to authentication. A number of attempts at providing better security for users may have good security characteristics in a theoretical environment, but fail when deployed in the wild or in tests with real users, often because usability has been given insufficient consideration[8]. Specifically, two factors are often ignored.

The first factor is that users are not good at paying attention to security indicators. They are particularly bad at noticing the absence of an indicator, or the absence of an indicator in a particular place (like the lock indicating an SSL-secured connection)[9]. Even when warnings become more obtrusive, such as changing border colors and pop-up boxes, users still ignore warnings and proceed to jeopardize their online security[10].

The second factor is that humans are susceptible to semantic attacks. Ideally, users should not be able to perform a "bad" action. If attackers can bypass the latest security efforts by altering their semantic attack approaches, then the value of those new security measures is limited to the change in difficulty of launching a semantic attack. Two real-world examples of this can be seen in solutions currently being deployed by U.S. banks for enhanced security. The first example is that of RSA SecurID tokens[11], which generate a random code every 60 seconds. RSA claims that this is an answer to phishing, but in reality semantic attacks are still very possible. There is nothing to prevent man in the middle attacks, provided that the information captured is used immediately.

Another example of security systems failing to secure against semantic attacks is Passmark's SiteSecure system. Passmark's SiteSecure system is a scheme to register a user's computer as a second factor for two-factor authentication [12]. The first time the user tries to log in from a given computer, they are asked 'secret questions' to establish their identity. Upon a successful response, a cookie is stored on the computer containing machine-specific information. This makes it more difficult for someone to access another's account, but the implementation contains severe vulnerabilities. The solution is vulnerable to man in the middle attacks that exploit the re-registration protocol [13][14]

IV. Concluding Remarks

We are looking at a wide range of areas that, when combined, may be promising for enhancing security on the web. It is unlikely that any solution will itself be a silver bullet; rather we believe that solutions will have to combine the approaches we have outlined in this paper. We cannot afford to be naïve, and must assume that the nature of phishing attacks is likely to evolve in the years to come, making new approaches, such as the semantic methods outlined previously, even more critical. Additionally, increasingly sophisticated and customized attacks might make the collaborative approaches more difficult to approach in isolation.

New authentication mechanisms will likely help, but seemingly all authentication protocols to date have had some weakness, which will inevitably be exploited by would-be attackers. We are in an ongoing battle, fighting a technological arms race, and we belive that the way to ultimately stay ahead in this battle is to combine what we have previously mentioned in a multi-pronged approach to thwart these attacks. Phishing is here to stay, and we cannot ignore the problem posed by it and other semantic attacks against security.

REFERENCES

[1] "Phishing activity trends report," Anti-Phishing Working Group, Tech. Rep., July 2005. [Online]. Available: http://antiphishing.org/APWG_Phishing_Activity_Report_Jul_05.pdf

[2] B. Schneier, "Semantic attacks: The third wave of network attacks," *Cryptogram Newsletter*, Oct. 2000. [Online]. Available: http://www.schneier.com/crypto-gram-0010.html#1

[3] C. Drake, J. Oliver, and E. J. Koontz, "Anatomy of a phishing email," in *Proceedings of the First Conference on Email and Anti-Spam (CEAS)*, 2004, available: http://www.ceas.cc/papers-2004/114.pdf. [Online]. Available: http://www.ceas.cc/papers-2004/114.pdf

[4] N. Chou, R. Ledesma, Y. Teraguchi, and J. C. Mitchell, "Client-side defense against web-based identity theft." in *NDSS*, 2004. [Online]. Available: http://www.isoc.org/isoc/conferences/ndss/04/proceedings/Papers/Chou.pdf

[5] "Netcraft toolbar," 2006. [Online]. Available: http://toolbar.netcraft.com/

[6] "Phishing activity trends report," Anti-Phishing Working Group, Tech. Rep., Nov. 2005. [Online]. Available: http://antiphishing.org/reports/apwg_report_Nov2005_FINAL.pdf

[7] F. Gandon and N. Sadeh, "Semantic web technologies to reconcile privacy and context awareness," *Web Semantics Journal*, vol. 1, no. 3, 2004.

[8] A. Whitten and J. D. Tygar, "Why Johnny can't encrypt: A usability evaluation of PGP 5.0," in *8th USENIX Security Symposium*, 1999. [Online]. Available: citeseer.ist.psu.edu/whitten99why.html

[9] R. Dhamija, J. D. Tygar, and M. Hearst, "Why phishing works," in *CHI (to appear)*, 2006.

[10] M. Wu, R. C. Miller, and S. L. Garfinkel, "Do security toolbars actually prevent phishing attacks?" in *CHI (to appear)*, 2006. [Online]. Available: http://www.simson.net/ref/2006/CHI-security-toolbar-final.pdf

[11] "Protecting against phishing by implementing strong two-factor authentication," 2004. [Online]. Available: http://www.antiphishing.org/sponsors_technical_papers/PHISH_WP_0904

[12] "Passmark security," Dec. 2005. [Online]. Available: http://www.passmarksecurity.com/

[13] T. Z. Bauknight, "Passmark's sitekey - answering the wrong question," 2005. [Online]. Available: http://www.cafeid.com/art-sitekey.shtml

[14] R. Dhamija and J. D. Tygar, "Phish and hips: Human interactive proofs to detect phishing attacks." in *HIP*, 2005, pp. 127–141. [Online]. Available: http://dx.doi.org/10.1007/11427896_9