Anti-Phishing Best Practices for ISPs and Mailbox Providers

A document jointly produced by the

Messaging Anti-Abuse Working Group (MAAWG)

and the

Anti-Phishing Working Group (APWG)

Version 1.01, July 2006

1. Introduction

The entire Internet community is familiar with spam attacks. Phishing[1] is a newer, related attack but one that results in severe privacy and security violations. Phishing can have serious negative financial ramifications for the individuals and organizations that are targeted. Phishing has become a major concern for ISPs, with pressure coming from both users who demand that service providers do more to protect them from attacks, and from the financial institutions targeted by these attacks.

ISPs are now forced to actively participate in globally reducing phishing attempts in order to mitigate customer churn and the potential for litigation. Unfortunately, there is a lack of public consensus as to how an ISP should best attempt this. This document describes some of the best practices used by members of the Messaging Anti-Abuse Working Group (MAAWG) (www.maawg.org) to combat phishing attacks.

2. The Nature of Phishing Attacks

Today the Internet community (and especially ISPs) generally have a reasonable understanding of the spam problem. In-house technology, third party solutions, and industry initiatives have given rise to a myriad of technologies to combat it. Phishing, on the other hand, is a relatively newer and more insidious threat. While spam and phishing are similar on the surface, phishing attacks are comparatively sophisticated and logistically different from spam in a number of ways. Understanding these differences is critical when considering a framework for protecting users from these attacks. This section examines the more important differences between spam and phishing.

- **Sophisticated**

The first important difference between phishing and spam is the messages and techniques used in phishing attacks are more sophisticated. Messages that phishers send out are usually carefully crafted to impersonate known, trustworthy financial institutions or organizations. Many anti-spam filtering systems are unable to distinguish phish e-mails from legitimate e-mails from these organizations. Phishers also systematically exploit software vulnerabilities in web browsers, web servers, and local operating systems in order to fool filtering or detection software, trick users, and steal as much information as possible. Their methods and techniques bear a much closer resemblance to “Black Hat” (destructive) hackers and virus authors than to traditional spammers.

- **Targeted**

Another difference between phishing and spam is that phishing messages are directed to a more targeted audience. Where spammers usually send to as many recipients as possible in an attempt to increase their response rate, phishers often target carefully chosen lists of e-mail addresses and

[1] Phishing is the practice of creating a replica of an existing web page to fool a user into submitting personal, financial, or password data. The word "phishing" comes from the analogy that Internet scammers are using e-mail lures to "fish" for passwords and financial data from the sea of Internet users. The term was coined around1996 by hackers who were stealing America Online (AOL) accounts by scamming passwords from unsuspecting AOL users. The first mention on the Internet of phishing was on the alt.2600 hacker newsgroup in January 1996; however, the term may have been used even earlier in the printed edition of the hacker newsletter "2600". (Source: Anti-Phishing Working Group (APWG) and WordSpy).

individuals. They are not just trying to improve response rates. More importantly, they are trying to evade less sophisticated data collection and detection systems (honey pots[2], for example) which might alert authorities to their attempted fraud.

- **Transient**

A third distinction between phishing and spam is the transience of phishing attacks. Phishing attacks are often very transient and short-lived, often occurring for only a few hours. In contrast, spam tends to be sent in frequent and large batches. Since phishing is a well-defined criminal activity, with constraints that are quite different from spam, transient attacks are necessary in order for the phisher to evade detection.

- **Dynamic**

The final important difference between phishing and spam is the dynamic nature of phishing. Phishing attacks, and the sites that host them, are dynamic, moving among servers very quickly. Whereas a spammer is advertising a product or service from a known website, a phisher is redirecting users to a private site designed to impersonate a financial institution or organization. Phishers typically exploit software vulnerabilities in web servers or server operating systems in order to install their own content. Many are sophisticated enough to be able to do this in an automated fashion, allowing them to compensate quickly and easily when a compromised site is discovered and taken offline. In fact, phishers cycle through compromised hosts quickly, regardless of their discovery, simply to confuse or obfuscate the true source of the attack.

3. Inbound Protection Schemes

3.1. Inbound Filtration of Phishing Messages

The most common front-line defense against phishing e-mails is the use of anti-phishing/anti-spam filtration technology at the outermost border Mail Transport Agent (MTA) or e-mail server. This is usually done using the same anti-spam software that the ISP already has in place to detect and filter spam.

Several techniques have been developed that are currently in use to filter spam. These include: IP address blacklists, Bayesian content filters, content heuristics engines, and content fingerprinting schemes, augmented by sender authentication. While all of these techniques are effective to varying degrees against spam, only some perform well against phishing. Here's a breakdown of what to expect with these different techniques:

3.1.1. Bayesian Filters

Bayesian classifiers filter spam based on their semantic difference from legitimate communications. Bayesian filters are in wide use, particularly in end-user anti-spam products.

[2] A honey pot is a computer system on the Internet that is expressly set up to attract and "trap" people who attempt to penetrate other people's computer systems. In the context of spam, honeypots are e-mail addresses that are created specifically for the purpose of collecting spam.

There are considerable and easily discernable differences between the content of spam messages and the content of legitimate messages. By contrast, phishing messages work to imitate legitimate messages in both content and form. Particularly, phishers will often copy real messages verbatim except for changing a couple of links to point to fraudulent sites. Most of the Bayesian filters in use today have been trained to detect spam. As such, it is rather difficult to detect phishing messages using Bayesian classifiers trained to detect spam. However, a Bayesian classifier specifically trained to detect phishing messages may do better. ISPs that deploy Bayesian filters should carefully measure the effectiveness of their filters against phishing messages.

3.1.2. IP Address Blacklists

The previous section described how phishers use compromised machines to host phishing web pages and to send out phishing e-mails. IP address-based (source-based) filters created to detect spam are particularly poor at detecting phishing messages because phishing messages often originate from otherwise "good" hosts. As with Bayesian filters, ISPs should carefully evaluate the efficacy of their source-based blacklist solutions for recognizing and filtering phishing messages.

3.1.3. Heuristics and Fingerprinting Schemes

Heuristics[3] and fingerprinting schemes[4] tend to perform reasonably well against phishing, especially if the solutions are specifically designed to detect phishing attacks. Heuristic solutions look for specific techniques used by phishers, such as encoding the name of a financial institution in the local part of a URL and using IP addresses as the host part of the URL. Fingerprinting schemes work by comparing known samples of phishing messages against incoming e-mail.

3.1.4. URL-based Filters

Some URL-based filters look for specific IP addresses, domains, or URLs where known phishing web pages are hosted. These IPs, domains and URLs are collected from reports of phishing messages gathered from e-mail users and "honey pots." As such, URL-based filters are fairly effective, but based as they are on limited reporting, can represent only a small sample of phishing activity at any given moment. Since phishers tend to frequently cycle through a large set of hosts, it is very difficult to have a comprehensive and updated list of bad IPs, URLs, and domains unless you have a high frequency of actionable reports. A few MAAWG members receive much higher rates of reports from their users due to custom software in the mail interface. These members have had great success in mining this data.

Recommendations:

1. **ISPs should conduct comprehensive field trials specifically targeted at phishing security before deploying an anti-phishing filtration solution. Ideally, ISPs should compare multiple solutions to determine their efficacy at stopping phishing attacks.**

[3] A rules-based problem-solving technique. Heuristic-based, anti-spam and anti-phishing filters look for telltale signs in e-mail messages that indicate the message is spam or phishing.

[4] Fingerprinting schemes compute a "hash" of a message's content that uniquely identifies the message. Fingerprinting schemes used in spam and phishing filtration products employ noise reduction techniques to generate fingerprints that don't vary with minor mutations of content.

3.2. Policy Considerations

In many messaging security systems, spam is often tagged, but then delivered to either a user's inbox or to a special "spam" folder, which allows users to review a message and personally determine whether it is spam or legitimate e-mail. With phishing messages, rather than delivering a detected phishing message to users, it is advised that the ISP drop the message or reject it at the SMTP level. Because phishing messages are designed to impersonate legitimate messages, many users cannot accurately assess the message as a phishing attempt. It is not at all uncommon for users to see a bank phishing message in their spam folder, assume that the filtering engine made a mistake, and click the link to the phisher's site.

Recommendations:

1. **Deny/reject phishing messages where possible.**

2. **When it is not possible to drop messages (due to user request, ISP policy, or legislative requirements), ISPs should indicate to the user that they are phishing messages and although they might look legitimate, they are dangerous and should be ignored.**

3.3. End-point or Client-Side Filtration

There are several free and commercial end-point security solutions on the market that plug in to users' e-mail software and filter phishing messages from incoming mail. In instances where an ISP is unable to provide server-level phishing filtration, these solutions can be effective. End-point solutions are also recommended so that users can be protected when they are accessing e-mail from multiple accounts, some of which may not reside on the ISP infrastructure.

Also, end-point security solutions are invoked when a user reads their e-mail, as opposed to server-side solutions that are invoked when the mail is delivered. Often, the latency between delivery and processing of mail is long enough for end-point filters to be updated, and hence, provide better security. The filtering latency is a good argument for providing two-tier phishing security schemes.

Recommendations:

1. **ISPs should encourage their users to employ end-point security solutions to combat phishing.**

3.4. Forgery Detection with Sender Authentication

E-mail authentication is becoming widely adopted. Among other things, e-mail authentication can be used to determine if the sender has forged the sender identity. Phishers often try to forge the information in the headers to make it appear as if the message originated from a legitimate institution. Sender authentication, where available, can often be used to detect this.

Recommendations:

1. **ISPs should filter or reject e-mail if they can unequivocally determine that the Sender's identity is forged.**

3.5. Hide Images from Untrusted Sources

Displaying images in "untrusted" e-mail messages puts recipients at great risk. Recently discovered security vulnerabilities in image display libraries underline the need to protect users, not only from subjectively offensive images such as pornography, but also from images that could abuse security breaches and install key loggers or other malicious software on the machines of unsuspecting users. E-mail providers long ago disabled JavaScript® and other executables for all incoming e-mail messages. There is now a positive trend to disable images by default, and to display images only when embedded in trusted messages.

Recommendations:

1. **ISPs should consider turning off images for all messages for which the identity and reputation of the sender cannot be established, and provide the user the ability to enable those images.**

3.6. Disable Hyperlinks from Untrusted Sources

Another method of alleviating the threat of phishing is to disable hyperlinks in untrusted e-mail. This makes it more difficult for phishers to trick users into clicking through to a fraudulent site.

Recommendations:

1. **ISPs should disable all hyperlinks in e-mail from untrusted sources.**

2. **ISPs should remove hyperlinks from suspected phishing e-mails.**

3.7. Visual Cues on Message Legitimacy

In parallel with their continued effort to block phishing messages, some mailbox providers also support a mechanism that conveys the authenticity of legitimate messages to their users within their e-mail interface. Although it has been suggested to allow senders to include images inside of the message itself, those images are easily faked. Instead, our recommendation is that these visual cues should appear in an area of the user interface (UI) that cannot be altered or spoofed. This may require changes to the e-mail client or web-based e-mail interface.

Recommendations:

1. **ISPs should consider providing their users visual cues (an icon in the message list view and/or in the message view) highlighting messages known to be legitimate and from trusted sources.**

2. **Each ISP should determine to what level they are comfortable endorsing presumed or known to be legitimate messages, and convey this endorsement level to users. The endorsement level depends on the ISPs' confidence in the underlying vetting technology and processes.**

3. **This recommendation applies to situations in which the ISP controls the user interface (webmail, proprietary MUA, plug-in for off-the-shelf MUA).**

4. Web Traffic Filtration

Phishing messages contain one or more links to a phishing website to collect user credentials. One way to render phishing attacks useless is to block access to these sites. There are several free and commercial efforts underway that provide lists of known phishing URLs to organizations that wish to limit access to these URLs.

Several browser tools that educate and empower users about the authenticity or fraudulence of websites they visit can effectively curb phishing victimization, even if users click on links within phishing e-mails. These tools are designed as browser plug-ins. They examine the links, as well as the content of the visited web pages, to make a decision regarding the safety of visited web pages. Browser-based, anti-phishing toolbars provide an additional "last line" of defense in preventing users from mistakenly providing credentials to a phisher.

Recommendations:

1. **Where possible, ISPs should enable short-lived blocks on confirmed phishing sites using firewalls and/or web-filtration products.**

2. **ISPs should bundle, distribute, or encourage their users to download web browser plug-ins that detect and restrict access to known or suspected phishing sites based on phishing URL feeds and/or predictive heuristic technologies. Certain plug-ins also authenticate legitimate websites and instill confidence in users about the safety of their web experience.**

5. Outbound Protection

Phishers often launch their attacks from compromised servers, without the knowledge of the owner of the server or the surrounding network. Phishers will either generate phishing e-mails directly from their own server or redirect messages through a botnet[5]: a large array of compromised machines under their control. In either case, the malicious traffic is transported by an unsuspecting carrier. Often, these unsuspecting carriers are end-user machines connected to an ISP via dial-up, DSL or cable modem. As a result, phishers often use the ISP e-mail infrastructure to send out phishing e-mails.

In some circumstances it may be within the capabilities of an ISP to attempt to filter outbound phishing attempts using anti-phishing filters. Several MAAWG members have reported great success in using their filtration solutions in "outbound mode" to stop phishing messages from leaving the ISP network. Another advantage of an outbound filter is that it might provide the ISP with a report of the location of phishing web pages. If these web pages are installed inside the ISP infrastructure, the ISP can decide to remove or restrict access to them.

Recommendations:

1. **ISPs should consider outbound content filters. When considering an inbound filter, ISPs should also evaluate the outbound capabilities of the solution.**

[5] Botnet is a jargon term for a collection of software robots, or bots, which run autonomously. A botnet's originator can control the group remotely, usually through a means such as Internet relay chat (IRC), and usually for nefarious purposes. (Source: Wikipedia.com)

6. Pharming and DNS Cache-poisoning Protection

Phishing attacks normally arrive via e-mail and request that a user visit a counterfeit site to enter personal information. Links to the phony site are always provided, and by looking at the URL address, it is often possible to determine if the site is indeed a fake. However, the successor to phishing attacks, called "pharming," is more difficult to spot. While a user may believe that they are visiting ebay.com, they may be unknowingly visiting a counterfeit site.

Pharming attacks work by attacking, or poisoning, the DNS systems used to translate Internet addresses. While attacks of this nature have been theorized for years, they have just recently been used for identity theft attacks. A pharmer, with access to a DNS server, will isolate specific sites, and route Internet traffic from the real site to a scam site. Unbeknownst to the user, he or she is now visiting a phony site, and any information entered could be used for malicious reasons. The unique danger of pharming attacks, as opposed to phishing attacks, is that a phony link does not need to be provided; even manually typing in a valid web address into a browser URL box can still lead to an attacker's site. Also, individual users do not need to be singled out, as any user of the compromised DNS server is vulnerable.

Recommendations:

1. **Currently, the best practice for avoiding pharming scams is to only access sites with certificates enabled. The ISP should educate their users about the dangers of pharming and encourage them to check for certificate validation when providing important credential information on a website.**

2. **ISPs should ensure that their DNS architecture is up-to-date. Old software or systems are more likely to be vulnerable to attacks, and can lead to the compromise of a DNS server, thus putting all users of that server at risk for pharming attacks.**

7. Phishing-related Customer Support Calls

Phishing problems inevitably generate support calls. Effective customer support processes and tools can save valuable time. A phishing or pharming support call is also a great opportunity for user education.

Recommendations:

1. **Remember that phishing and spam are not synonymous. Train your support representatives to recognize the difference.**

2. **If a user reports suspicious e-mail asking for personal information, the ISP should inform the user of the dangers of phishing attacks, and warn him or her against giving out personal information online. The user should be further advised to send a copy of the e-mail to the ISP, so it can be used to update filters.**

3. **If the user believes that he or she has been scammed, the user should be urged to file a complaint with the appropriate anti-fraud organization such as the Federal Trade Commission (FTC). APWG maintains a list of anti-fraud organizations at http://antiphishing.org/resources.html#antifraud**

4. **Customer support processes should be in place for quick remediation in cases where a suspected phishing e-mail or site is sent from or hosted by the ISP itself.**

5. **Customer support should also direct users to consumer education resources that enables them to understand the nature and scope of these threats, and which describe measures the ISP is taking to protect users.**

8. ISP-to-Phishing Target Communications

Where possible, ISPs should try to communicate early knowledge of phishing attacks to the targeted institution.

Recommendations:

1. **ISPs should communicate knowledge of phishing attacks to the targeted institution via the Anti-Phishing Working Group at www.antiphishing.org, or via a similar, regional organization.**