Phishing Attacks and Web Spoofing

Tarjei K. Mandt and Håvard Hasli

Last revision; March 29, 2006

Abstract

Phishing attacks and web spoofing are on the rise with an increasing number of reported incidents. Hackers and computer criminals resort to social engineering to fraudulently acquire sensitive information, such as credit card details, passwords, and social security numbers. The targeted victim is given the impression that the attacker is a trustworthy person or business, and confidently hands over the requested information. In this article we will look at the various methods, some of which are commonly used today, and show how to detect and counter them.

1 Introduction

Identity theft is as old as the pen and paper itself. Throughout history, it has been a popular way for criminals of acquiring sensitive information for financial gain. The term *phishing* originated in the way people "fish" for information, by persuading individuals into giving away valuable information, often through popular Internet communication channels such as e-mail or instant messaging. Phishing is, however, not limited to the electronic world. On numerous occasions, devices attached to automated teller machines have been found that record the user's credit card information. When a credit card is inserted, the magnetic card reader attached to the real ATM card slot records the credit card information.

According to Symantec Security Response in a recent article [IT Pro], there are 33 million unique phishing messages every week. Estimates indicate that the fraud incurred by phishing total over 1 billion USD each year. Even more disturbing is it that studies made by the Anti-Phishing Working Group [APWG] conclude that phishers succeed with as much as 5 percent of all recipients. Because of this, companies fear that customers eventually will lose confidence in e-commerce.

Phishing methods used today are a cross between social engineering and spoofing techniques, in which users are made to believe they deal with a trustworthy person or business. *Web spoofing*, often used together with e-mails, is perhaps the most frequently used phishing attack. A trusted web site is reproduced by the attacker in a way that makes it hard to separate from the original. The traditional way of doing this is by sending the victim an email and request the login information.

The e-mail is as with the web site, made to look exactly what the targeted company would send, using various spoofing methods. Conveniently enough, the e-mail also provides a link (the attacker's site) to which the victim may proceed. Usually the URL will give you a clue about its legitimacy. It is, however, not always so easy to detect these scams. As we will see in the following chapters, far greater and more sophisticated methods may be used for successfully launching web spoofing attacks.

2 Phishing Attacks

In this section we will provide an overview of the various phishing attacks that are used today.

2.1 Spoofing e-mails and instant messages

There are several different types of phishing attacks which can be divided into different categories. The oldest form that dates back to the mid-90ies is known as the *e-mail spoofing attack*, in which spoofed e-mails were used to persuade the victim into sending passwords and account information using electronic mail. This method is, however, not as much used anymore as over the years, users have become more careful in sending sensitive information unprotected over e-mail. This is partially because banks and online services now only handle such information directly from the web site, using proper encryption such as SSL.

Today, such attacks are carried out in combination with web site spoofing methods, which is probably the most commonly used phishing technique to date. Attackers target global organizations and companies such as eBay and PayPal to reach out to millions of potential victims. The e-mails sent out are made to look authentic, including company logos and guidelines stating that the account information should not be transferred to any third party, as a legitimate business also would stress. E-mails today are also easily forged due to the lack of proper authentication in the aging SMTP protocol. If a reply is irrelevant, an attacker can replace the `reply-to` field in the

e-mail header with a more fitting address such as `service@paypal.com` or `sales@ebay.com`.

Another way of spoofing e-mails and directing users to malicious web sites can be done by masking the link URL. In HTML this is usually done by using the legitimate url as the link text, while the link itself point to the malicious web site. An example of this would be:

```
<a href="http://paypal.com.login-user1414.info">
  https://www.paypal.com/cgi-bin/webscr?cmd=login
</a>
```

Inexperienced users may overlook the actual URL, which in the Windows environment can be revealed by positioning the mouse cursor above the link.

Instant messaging services has also in recent years been targeted by phishers. According to a recent article [Leavitt 2005], more than 506 million people worldwide will use an IM product by 2008. Further estimates indicate that there will be as many as 78 million enterprise users, as IM has become popular for businesses for its real-time capabilities. Phishers may use IM to spread messages containing links to counterfeit web sites, and retrieve information such as user ID and password. Furthermore, IM worms can let an attacker hijack and send messages with phishing-related hyperlinks from victims' clients to their IM contacts. This could make the receiver believe that a person they know sent the message and that clicking on the hyperlink is safe.

2.2 Cross-site scripting attacks

Cross-site scripting attacks (XSS) are carried out by using code injection or custom URLs in a valid web-based application. Scripts vulnerable to XSS attacks fail to properly sanitize user-supplied input and are usually the result of poorly developed web applications. Typical ways of injecting XSS into valid URLs include:

1. Full HTML substitution
   `index.php?url=[XSS]`
2. Inline embedding of script
   `index.php?client=<SCRIPT>[XSS]...`
3. External script code
   `index.php?post=[XSS]..phisher.com%21code.js`

XSS attacks can benefit phishers in various ways. Using HTML client-side scripting it is possible to embed the desired code or information into a trusted web site. One way of doing this is by entering code into a web form, such as a bulletin board message, which then is rendered in a different way than intended. In the example above, `url` can define the address of an inline login page. Without proper checks in index.php, an attacker can inject the url to a site he or she controls. While the injected url may reveal that a phishing attempt is taking place, the attacker can easily obfuscate it using different encoding schemes, such as escape encoding and UTF-8 encoding.

The following example can be used to demonstrate a XSS attack:

1. Alice trusts example.com to run JavaScript on her machine.
2. Bob injects a malicious script (for example using a bulletin board) that asks for people's credit card numbers and makes these available to him.
3. Alice visits example.com and Bob's script ask for her credit card number. Alice trusts example.com and provides her credit card number.
4. Bob has now stolen Alice's credit card number using XSS.

2.3 Session attacks

As HTTP and HTTPS are stateless protocols, web servers need to use custom methods for keeping track of users and authentication. Therefore, many systems use what's commonly known as a session identifier to keep track of its users. The SessionID can be stored in hidden fields within a web page, in a cookie, or even embedded in the URL itself. If such a system is poorly implemented, it may be vulnerable to phishing attacks.

A potential vulnerability may exist in allowing preset SessionIDs, and pre-defining them upon accessing a web site. For example, the attacker can send out thousands of e-mails, and provide a link with a pre-defined SessionID to a bank some might use. The attacker then polls the bank web site using the same SessionID until someone is authenticated. Once this happens, the server will allow anyone who uses the preset SessionID, including the attacker.

The following URL is an example of a web page that uses SessionID.

```
http://www.onlinebank.com/login.php?sid=241ae2k1m95fl
```

2.4 Key logging

Another common method used in phishing is *key logging*, in which a piece of software is installed on a computer to monitor keys that are pressed.

This way, attackers can capture login sequences and grab usernames and passwords. Key loggers may be disguised as trojans and can be difficult for the novice user to identify.

3 Web spoofing

In *web spoofing* the attacker creates a copy of the web site from which phishing is performed. The copy is identical to the real web site and runs on a server that the attacker controls. Web spoofing is really about making the copy convincing enough to make victims believe they are in good hands. To achieve this, various methods are used.

3.1 Name similarity

It is not unusual that people mistype the name of the web site they are looking for. Because of this, it is possible for attackers to set up spoofed web sites on similar domain names. Usually these sites are only in it for the potential traffic gained, as they often redirect to sites that sell products or services and might not relate at all to the site spoofed. However, this doesn't have to be the case. An attacker could easily set up a spoofed site under a similar domain name, and catch unaware users' login information. As many companies and organizations register similar domain names to counter act this possibility (for instance, try mistyping *google.com*), some may also think that a spoofed website is legitimate.

There are numerous examples of sites that use name similarity. These include `whitehouse.com`, `ebau.com`, and `nasa.com` just to name a few.

3.2 URL spoofing

URL spoofing deals with the different ways of making a spoofed site URL resemble a genuine site URL. In doing so, the attacker may have a better chance at succeeding, especially with inexperienced users who are unfamiliar with phishing. In the code segment mentioned in section 2.1, the domain name *login-user1414.info* has a subdomain by the name *paypal.com*. It is not unlikely that some users only pay attention to the first part of the FQDN[1], *paypal.com.login-user1414.info*.

Another way of masking the URL is done by including a user name and password. Web servers that require authentication may be accessed using the

[1]Fully qualified domain name - consists of a host and domain name, including top-level domain.

URL string format *username:password@domain.com*. User name and passwords in URLs may be used regardless of whether the web server enforces this or not. The information is simply ignored if not. User names are not limited to just letters and numbers, so for instance *www.paypal.com* could very much be a valid choice. Consequently, an attacker could construct a URL such as `http://www.paypal.com:80@192.168.0.1/` where *www.paypal.com* is the user name, *80* is the password, and *192.168.0.1* is the malicious site. It is also possible to omit the password completely. The method is, however, not as much used anymore as browsers now notify when a user name and password in the URL is used (and that a phishing attempt could take place).

3.3 IDN spoofing

Since the introduction of *internationalized domain names* (IDN), domain names may now also include country-specific characters. Unfortunately, some foreign language characters look almost the same as certain Latin characters, and may therefore also be used in phishing attempts. Not only does this allow attackers to register domain names that look exactly like another, but it also allows the use of security certificates which appears to be for a legitimate domain. A good example is the paypal.com case in which the Cyrillic *a* replaces the Latin *a*, `http://www.p&#1072;ypal.com/s`. In Unicode, decimal 1072 represents the Cyrillic *a*.

For the Unicode strings to be mapped into the limited character set supported by the DNS (domain name system), *Punycode* is used. Punycode is applied to each component of a domain name address (subdomain, domain name, and top-level domain) which contains characters not found in the ASCII character set. For each translated Punycode string, a prefix `xn--` is added. Any foreign character is then stripped and replaced by a trailing code. Using the same example as above, the result would become `www.xn--pypal-4ve.com`.

Because Punycode enables websites to use full Unicode names, web browsers including *Firefox* and *Opera* now use a white-list of TLDs[2] that have policies for which characters are permitted and procedures for making sure that no homographic domains are registered to two different entities. While a white-listed TLD will be displayed in Unicode, any untrusted TLD will be displayed in Punycode with the `xn--` prefix. Dot com is not part of this list and will therefore be displayed in Punycode.

[2]Top-level domain - the last part of the Internet domain name such as com, net, or org

3.4 DNS spoofing

DNS spoofing attacks or *poisoning* attacks are attacks in which attackers attempt to feed incorrect mappings between IP addresses and host names to the DNS server. As DNS queries are usually submitted over UDP, servers cannot rely on the transport protocol to maintain state of the DNS connection. Therefore, in order to match a response with a query, DNS servers include a numeric query ID in the DNS payload. If the attacker can predict the query ID, it is possible to craft a spoofed response before a real response is returned to the DNS server. The DNS usually believes the first response it receives, and discards any additional responses which then are considered duplicates. Consequently, anyone who looks up the spoofed domain record will be redirected to the attacker's site.

Another way of performing a DNS cache poisoning attack, can be done on the victim's computer. Every system has a host file in its system directory used to associate host names with IP addresses. This is actually the job of a DNS server, but by adding records to the hosts file, one may hard code domain name translations and redirect users to different sites.

The hosts file is located in `%SystemRoot%\system32\drivers\etc` in the Windows environment, and may also be found under `/etc` in UNIX-based systems. Each line in the hosts file represents an entry. The first column specifies the IP address followed by the corresponding host name. Most systems map `localhost` to the loopback address as shown below.

```
127.0.0.1       localhost
```

Normally, when you attempt to access `domain.com`, a request is sent to a DNS server the find out the IP address for that domain name. Once this has been done, the HTTP request is forwarded to the proper web server. However, if we were to insert a custom entry for `domain.com` in the hosts file, the request would be forwarded to this address instead.

```
127.0.0.1       localhost
192.168.0.2     domain.com
```

An attacker could use this method to direct users to a web site that he or she controls, even if the victim types `http://domain.com` in the address bar of the web browser.

3.5 Proxy spoofing

It is also possible to redirect users to malicious sites by defining proxies in the browser configuration. This is usually done by having the user install

some sort of web extension (aka trojan/spyware) which then can override the settings present in the web browser.

4 Anti-Phishing

In this section we provide means to counter phishing, focusing on what users can do, what software that can be used and what laws that is in place for handling this type of crime.

4.1 Users

As phishing has been a part of our internet life since the 90s, more and more people get aware of, (or get a sense of), that something is wrong with their email or the site their currently viewing. Still, there will always be newcomers to the "net" and some training is required for people to understand the difference of "good and bad" email/sites.

One of the most common ways to fool a user is focused toward the user to reveal its username, passwords or account information. Most companies will **never** ask for your information in a mail and in case they loose your password or account due to an error, they will simply make you a new one.

By comparing old mail, that you know is from the company, and the new one you just got, you may see tiny differences and elements out of order, i.e. the header of the mail don't contain the usual greeting with your customer number.

Spear attack is a recently new phising attack focused toward a specific company. This form of attack has been adopted by several companies to test their customers/employees to see how they respond to false mail/sites. `http://www.antiphishing.org/consumer_recs.html` includes a good list of how to behave when unsure about whether or not you are facing a phishing attempt.

4.2 Software

While training is not always an option there has been developed different software to help you combat the problem. In a Pc World article [PCW] Tom Spring has tested three different tools, that use different approaches to detect phishing. The first one is *Anti-fraud toolbar* from Cloudmark[3] is based on a blacklist of known "bad" sites. When a user attempts to enter a site on the

[3] `http://www.cloudmark.com/`

blacklist the program will issue a warning stating that the site is not safe to go to.

The second one is *Trustbar* from Comodo[4] and it is based on a white-list of "good" sites. When you surf on a site that is on the white-list the companies name will appear on the toolbar, when not on the white list a warning will be issued, but it doesn't prevent you from surfing further.

The last one reviewed was *Clear search anti phishing* from Phishing.net[5] which uses 35 different properties of a phishing scheme to identify such attempt. In his tests the solution had problems identifying some of his known phishing email and some other known bogus sites. The company claimed it was a malfunction in the software he used and that it was a unique case.

Tom Spring found that none of the software was in any way complete and could find all attempts on phishing, but it seemed that the blacklist solution worked best against his test, with white-list at the second place.

A list of other tools can be obtained at `http://www.antiphishing.org/solutions.html\#toolbars`

4.3 Laws

The Anti-Phishing Working Group [APWG] reports that the number of new phishing messages climbed at a monthly rate of 38 percent in the last six months of 2004. The number of new phishing websites has climbed 24 percent per month since last August.

Until now there has been no really good law that takes care of the specific act of phishing. Most of the law has focused on dealing with the perpetrators after the crime has been committed, but as of March 1 - 2005 this will change.
From Pc World March 7 - 2005 [PCWB]:

> WASHINGTON – A bill introduced in the U.S. Senate last week would allow prison time of up to five years and fines of up to $250,000 for people who design fake Web sites for the purposes of stealing money or credit card numbers. The Anti-Phishing Act of 2005, introduced by Senator Patrick Leahy (D-Vermont), would outlaw "phishing", in which scam artists design Web sites to look like real banking or e-commerce sites, then e-mail spam to people saying they need to re-enter their account or credit card numbers at the bogus site.

[4] `http://www.comodogroup.com/`
[5] `http://www.phishing.net/`

This law is mainly focused toward USA but several European countries have followed the example and tighten their law in these areas. Such laws should make it more unattractive for people to try phishing, but there would still be a few organized people willing to take the risk.

5 Conclusion

Phishing and web spoofing has become a serious issue to the global internet community. More organizations provide greater online access for their customers, and rely heavily on online services in the way they do business. This dependancy open up new ways for professional criminals to use phishing techniques to steal personal financial information.

While the threat is considered serious, there are plenty of ways for businesses and individuals to counter it. Understanding the tools and methods the phishers have at their disposal is a crucial element in defending against these attacks.

References

[Kirda, Kruegel 2005] "Protecting Users Against Phishing Attacks with AntiPhish", Engin Kirda and Christopher Kruegel, COMPSAC 2005.

[IT Pro] "Spike in Phishing and Malware a Danger to IT", IT Pro 2005.

[Ollmann 2005] "The Phishing Guide: Understanding and Preventing Phishing Attacks", Gunter Ollmann, NGSSoftware Insight Security Resarch 2004.

[APWG] "The Anti-Phishing Working Group", http://www.antiphishing.org/ visited on 2005-11-16.

[INPS] "Inside Network Perimeter Security", First Ed. 2003, New Riders Publishing.

[Leavitt 2005] "Instant Messaging: A New Target for Hackers", Neal Leavitt, IEEE Computer Society 2005.

[PCW] "Fight Fraud and Phishing With New Tools",Tom Springer, http://www.pcworld.com/reviews/article/0,aid,120501,00.asp visited on 2005-11-20.

[PCWB] "Proposed Law Aims to Fight Phishing",Grant Gross, IDG News Service Monday, March 07, 2005 http://www.pcworld.com/news/article/0,aid,119912,00.asp visited on 2005-11-28