CERIAS Tech Report 2015-10

Website Forgery: Understanding Phishing Attacks & Nontechnical Countermeasures for Ordinary Users

by Ibrahim Waziri Jr
Center for Education and Research
Information Assurance and Security
Purdue University, West Lafayette, IN 47907-2086

Website Forgery: Understanding Phishing Attacks & Nontechnical Countermeasures for Ordinary Users

Ibrahim Waziri Jr.
PhD Candidate in Information Security
CERIAS
Purdue University
West Lafayette, IN, USA
iwaziri@purdue.edu

ABSTRACT

Website Forgery is a type of web based attack where the phisher builds a website that is completely independent or a replica of a legitimate website, with the goal of deceiving a user by extracting information that could be used to defraud or launch other attacks upon the victim. In this paper we attempt to identify the different types of website forgery phishing attacks and non-technical countermeasure that could be used by users, (mostly by non IT users) that lack the understanding of how phishing attack works and how they can prevent themselves from these criminals.

Keywords

Phishing, Information security, Website forgery.

1. INTRODUCTION

In this technological era, everyone connects to the internet either using a computer or some sort of a mobile device. Financial transactions, academic registrations etc. are mostly conducted online. However, the percentage of non-IT users using these services outweighs that of IT aware users. According to the Internet World Stats, as of June 2014. "There are 3.03billion internet users out of the 7.1billion population in the world"[1]. A study by Gartner [2] shows that an "estimated 57 million American adults received e-mail attacks from phishers (hackers) who pretends to be trusted service providers to steal consumer account information. From the survey Gartner concludes that more than 30 million people were absolutely sure they were victims of a phishing attack". Therefore, if more than 50% of American internet users are victims of phishing attacks (considering the US is one of the top IT countries in the world). We can assume that at least more than 50% of the 3.03billion internet users around the globe are vulnerable to phishing attacks. A more recent report by RSA recorded an estimated losses of over USD $5.9billion in 2013 globally. With U.S being the most targeted country suffering over 60% of worldwide phishing volumes [3]. The volume of phishing attack increases year over year. While cyber threats (like phishing attack) are getting considerable attention from the users and media. We are still yet to see a decline in the volume of threats every day. The majority of phishing attacks victims are users that lack awareness of how phishing attack works. Later in this paper we will define what phishing attack is, how phishers implement phishing attacks and how users can differentiate between a legitimate website and a malicious one.

2. RELATED WORK

Phishing is a major threat to the cyber world, as a result numerous works about phishing prevention has been conducted, and however none of the work is targeted towards non-technical users.

Some works introduce tools that model and describe phishing attacks, allowing visualizations and quantification of the threat on a given complex of web services [4]. The authors use a new model to describe some new phishing attacks, some of which belong to a new class of abuse called *context aware phishing attacks*. The authors describe ways of using the model introduced, to quantify the risks of an attack by means of economic analysis and methods for defending against the attacks. The first part of the paper is a theoretical applicable model covering a large set of phishing attacks aimed towards developing and understanding threats relating to phishing. The second part of the paper is the description of what is *a context aware phishing attack* - which is defined as a threatening attack that is likely to be successful not only against the most gullible computer users. The authors claimed *context aware phishing attack* is mounted using the messages that somehow from their context are expected or even welcomed by the victim. The last part of the paper is a discussion of how to address the threats described both in specific and generic shapes.

Another work [5] focuses on identifying several of the technical capabilities that are used to conduct phishing scams, reviewed the trends and provided countermeasures. This study has similar goal with our study, however this study focuses on the technical aspect of phishing and target the IT aware users. The study identifies the tools used in implementing and delivering phishing attacks; these tools as outlined in the study are: Bots/Botnets, Phishing Kits, Technical Deceit, Session Hijacking, Abuse of Domain Name Service (DNS) and Specialized Malware. The paper also provided some countermeasure which includes; Awareness and Education, Strong Authentication and

Authorization, and the last Virus, Spyware and Spam Prevention.

The study "Protecting Users against Phishing Attacks with AntiPhish" [6] presents a browser extension "AntiPhish" which aims to protect users against spoofed web site-based phishing attacks. AntiPhish tracks the sensitive information of a user and generates warnings whenever the user attempts to give away this information to a website that is considered malicious. AntiPhish is based on the premise that for inexperienced, technically unsophisticated users, it is better for the application to check the trustworthiness of a web site on behalf of the user. The study claimed that AntiPhish cannot be fooled by obfuscation tricks such as similar sounding domain name.

Another study [7] proposes a new end-host based anti-phishing algorithm called "LinkGuard", by utilizing the generic characteristics of hyperlinks in phishing attacks. The characteristics are derived by analyzing the phishing data archive provided by the Anti-Phishing Working Group (APWG). The paper claims that LinkGuard can detect not only known but also unknown phishing attacks. LinkGuard was implemented in Windows XP. And the experiment result shows that LinkGuard is effective to detect and prevent both known and unknown phishing attacks with minimal false negatives. LinkGuard detected 195 out of the 203 phishing attacks.

A similar tool "PhishNet" is also introduced in a study [8]. PhishNet starts with observation that attackers often employ, such as simple modifications like changing top level domain to URLs. PhishNet exploits this observation using two components. The first component proposes five heuristics to enumerate simple combinations of known phishing sites to discover new phishing URLs. The second component consists of an approximate matching algorithm that dissects a URL into multiple components that are matched individually against entries in the blacklist. In the evaluation with real-time blacklist feeds, PhishNet discovered around 18,000 new phishing URLs from a set of 6,000 new blacklist entries.

Another study analyze empirical data on actual phishing website removal times, and the number of visitors that the websites attract, and conclude that website removal is part of the answer to phishing, but it is not fast enough to completely mitigate the problem [9]. The study identifies a subset of phishing websites (operated by the 'rock-phish' gang) which through architectural innovations have extended the average lifetime of their phishing websites.

Website Forgery Prevention is a paper that presents phishing prevention approach based on mutual authentication [10]. Authentication process changes so that the user is obligated to interact in a new authentication step which provides the authenticity of a website. Another paper [11] proposes a new class of Human Interactive Proofs (HIPs) that allow a human to distinguish one computer from another. Unlike traditional HIPs where the computer issues a challenge to a user. This type of HIP can be used to detect phishing attacks, whereby websites are spoofed in order to trick users into revealing private information.

A paper "Why Phishing Works" [12] claims to provide the first empirical evidence about which malicious strategies are successful at deceiving general users. First, it analyzed a large set of captured phishing attacks and developed a set of hypotheses about why these strategies might work. Then, assessed the hypothesis with a usability study in which 22 participants were shown 20 websites and asked to determine which ones were fraudulent. The result shows that 23% of the participants did not look at browser-based cues such as the address bar, status bar and the security indicators, leading to incorrect choices 40% of the time. The result also found that some visual deception attacks can fool even the most sophisticated users. The results illustrate that standard security indicators are not effective for a substantial fraction of users, and suggest that alternative approaches are needed.

A questioner study was conducted to find out if security toolbars actually prevent phishing attacks [13]. Because toolbars are designed for humans to use, the study attempts to evaluate the toolbars for usability. The study conducted a two user studies of three security toolbars and other browser security indicators, and found them all ineffective at preventing phishing attacks. The study claims that even though subjects were asked to pay attention to the toolbar, many failed to look at it; others disregarded away the toolbars warning if the content of web pages looked legitimate. The paper states that "many subjects do not understand phishing attacks or realize how sophisticated such attacks can be".

From all these studies, we can see how phishing attacks poses threats to user information and visibility online. New studies based on new threats are conducted every day to tackle the challenging phishing threats.

3. THE LOGISTICS OF PHISHING ATTACKS

Phishing itself is not a new concept, but as a result of online services, it is increasingly used by phishers to steal user information and perform identity crimes in recent years. To carry out phishing scams, phisher's uses social engineering techniques which involve human interaction and fooling people into revealing information that is a treat to them. Another form of attack phishers use is sending spam emails that include links (URL's) to malicious websites. Another example of a technical from of phishing is the use of malwares where the phisher creates malicious software designed to capture user information. This form of attack is a technical one, in which the phisher needs to have some technical knowledge of how to code (ability to design

software). And then the software will be downloaded to the targets devices. Users tend to download these software's (malware) without prior knowledge of the malwares existence. The top identified 5 ways to get infected with malware are visiting adult websites, installing unpatched OS & programs, (cracks, keygens, serials etc.), peer-to-peer sharing (torrents), advertisements (pop-ups) and untrusted links in emails [14]. For website phishing attack to take place we have three players in place: The Attacker, The Compromised Server and The Victim.

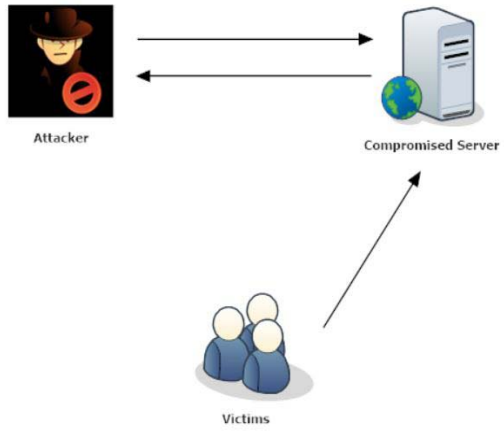

Figure 1: Phishing Players

The attacker is the phisher whose goal is to extract information. The compromised server is where the malicious website is hosted and along with some database for storage. The victim is always the innocent user.

4. HOSTING A WEBSITE PHISHING ATTACK

Phishers use different approach to implement phishing, but all website phishing attacks are hosted using either free webhosting, paid webhosting, website hijack or device hijack.

4.1 Free Web Hosting

In free webhosting phishing attack, the phisher utilizes the advantage of building a free website by using one of the free webhosting provides (wordpress.com, wix.com etc.). However with the services being free, there are certain limitation to these free webhosting sites with the most common being;

i. The user cannot have a domain name, only a subdomain, with the domain name being the free webhost provider.

ii. Ads and other pop ups might appear on the webpage, because the services are free. The webhost provider usually shows ads in other to make money.

Other limitations like the number of webpage allowed, and content size exists. But for most website phishing attack, all which is required is just a single webpage that could be used to deceive the victim. An example of a free hosting URL is https://justanexample.freehosting.com where freehosting.com is the free hosting company's domain name. And justanexample is the subdomain acquired by the user for free.

4.2 Paid Web Hosting

In the case of paid webhosting, the phisher pays a hosting company to have a website and a domain name. This usually tends to be more legitimate looking website, however with little investigation, a user can identify if the website is what it claims to be or not. In the case of replica website, typo squatting is common in the domain name. (Mostly with banks, the domain usually looks like that of the legitimate bank's website. With a little typo squat that require attention to be identified). Most paid phishing webhost are purchased for a short period of time, 1 or 2 years. More identification techniques will be discussed. An example of a paid webhosting domain is https://justanexample.com where justanexample.com is a paid domain name.

4.3 Website Hijack

Website hijack is another commonly means of hosting a phishing website. The phisher usually finds a way to infiltrate a legitimate website and then create another page within that website and use it to implement website phishing without the awareness of the legitimate website owner. In these cases, the malicious page is always contained in multiple folders within the domain name.

4.4 Device Hijack

This is a bit more technical, because the phisher doesn't only hijack a website or create a malicious webpage. Instead the phisher hijacks a whole webserver or another device and turns it into a webserver allowing the phisher to create and have access to all websites hosted on the device.

5. CHRONOLOGY OF WEBSITE PHISHING ATTACK

Figure 2 depicts the chronology of a website phishing attack. Usually the phisher starts by deciding on which of the website hosting method will be used to build the malicious website. Once that has been decided, the attacker then builds the malicious website and connects the website to a database that could be used to store stolen information.

Next, the phisher then sends out an email containing a link with the URL of the malicious website. The victim then checks the email and visit the malicious website using the URL in the email. Within the malicious website, the victim not knowing the difference between a malicious and a

legitimate website then uses his/her credentials on the website. The credentials then get stored automatically in the database.

The final step is when the phisher retrieves the stored credentials from the database.

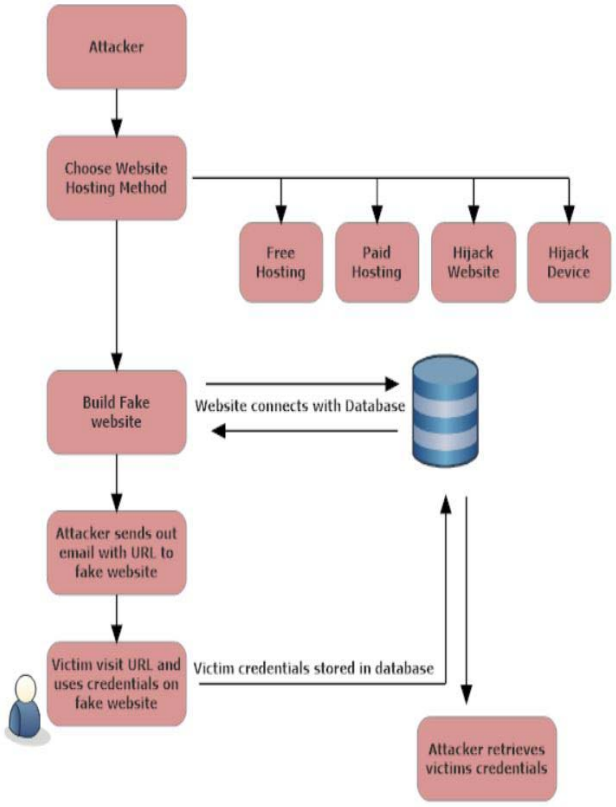

Figure 2: Website Phishing Attack Chronology

6. IDENTIFICATION & COUNTERMEASURES

Different ways and methods are available to verify if a website is legitimate or malicious. Here we will attempt to identify the common ways of website verification.

6.1 URL Obfuscation

URL Obfuscation misleads the victims into thinking that a link or web site displayed in their web browser is that of a trusted site. Checking the URL is one of the most important steps in identifying a malicious website. The phisher can always make a replica of a certain page, but never the URL. A website can have multiple domain names, but a domain name can only be assigned to one website. Therefore it is very important to make sure that the URL is that of the legitimate website.

https://justanexample.freehosting.com:87/free.money/index.html

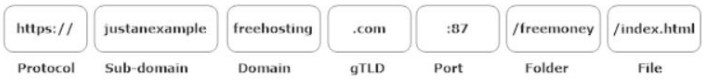

Figure 3: Understanding URLs

6.2 URL Redirection (Hyperlink)

One of the simplest techniques for hiding the actual destination of a URL is the use of hyperlink and redirection. The use of a legitimate URL within an anchor element but having its href attributes point to a malicious site. Basically, hiding a malicious website link underneath a known legitimate URL, thus clicking on the legitimate looking URL sends the user to a malicious site.

Countermeasure: URL information is typically displayed in the web browsers status bar. And therefore this deception can be detected by examining the link in the status bar to see if it corresponds to the claimed link.

6.3 Use of Images

It is now common for an email to contain images and phishers are taking advantage of this by constructing scam email that contain images and look real like that of legitimate websites. Most times they even use the original images and use it as a hyperlink like that of a URL redirection.

Countermeasure: Same method used to detect URL Redirection can be applied to use of images deception and hovering the cursor over the image to see the redirection link to examine the address bar

6.4 Domain Typo Squatting

Domain Typo squatting is a form of phishing attack domain name registration in which the phisher registers a domain name that is similar to the legitimate website domain name. Typo squatting is very common with paid hosting (applies to free hosting too) considering the phisher cannot use the legitimate website domain name, so a similar name with a little typo squatting that is difficult to be identified is then registered as the malicious domain name An example of this is:

Legitimate URL: http://mybank_example.com

Malicious URL: http://mybank~example.com

Notice how the Legitimate URL has a "_" and the malicious uses "~" instead. These common errors are hardly noticed by users.

Countermeasure: A good way of verifying typo squatting domain names is by checking the WHOIS information of the domain. Checking WHOIS involves querying domain names, a lot of services provide the identity and information about domain owners. Details such as the date of website

creation, the expiry date (a legitimate bank or merchant will never have a 1 or 2 year domain subscription). There are various WHOIS providers, few are: *whois.net, whois.com, centralops.net* among many others. And another way of easy verification is by using "Google" to search the legitimate bank or merchant name. (Trusted domains always appear at the first page of Google search) and corresponding it to the claimed URL.

6.5 Use of Free Domains

As stated earlier, because most hosting companies provide free hosting service to attract customers but with limitations. Some of which include:

- Not having a domain name, only a sub-domain.
- Having ads on the webpage.
- Number of allowed pages (most time 1 page limitation)

Most phishers take advantage of using free hosting companies to host malicious websites and then get a sub-domain without having to pay for the services of getting a domain name. The reason they usually do that is because they know people don't really pay attention to the address bar and URL's. Most people pay attention to the content of the website. And if it looks real then it's assumed safe.

URL with subdomain:
http://mybankexample.freehostingcompany.com

URL without subdomain:
http://mybankexample.com

In a free hosting service, the URL always contains the name of the free hosting company as indicated above (freehostingcompany.com).

Countermeasure: As always, paying attention to the URL and address bar plays an important role in differentiating legitimate and malicious websites. With the use of free domains, it is always a good practice to cut the URL to the domain name. That is removing the sub-domain and visiting the domain name website to see if it is what it claims to be. In most cases the domain name is a free hosting company. A legitimate business website will never use a free hosting service to conduct business transactions. From the example of subdomain URL above, cutting back to the domain means removing the subdomain "mybankexample" from the URL and visiting "freehostingcompany.com". (A subdomain name is always separated by a *period* from the domain name, and the domain name is the name closest to the TLD ".com")

These are some of the commonly identified website forgery phishing attack as of today, however new method and approaches are always implanted by phishers with no prudent. Other commonly practiced countermeasure exists and it is advised to be practiced on every website before sharing any information. Some of these countermeasures are:

- Awareness and Education about the usage of internet services
- Paying attention to web browser, tool bars and address bars
- Ensuring that financial and information sharing websites are secured with the use of https protocol and SSL certificates. (SSL certificate ensures that all information used on that specific website is encrypted, and it also uses the https protocol, which is the secured version of http). To ensure that a website is secured, always check for the padlock sign close to the URL and https in the URL. Clicking on the padlock sign reveals the SSL certificate. Phishers can't get an SSL certificate for malicious websites because it involves providing ones identity and that is something they can't afford to. (Remember they live in shadows).
- Authentication and Authorization – Usage of two factor authentication in all information related websites. Usually it involves the use of authentication (providing password or PIN) and then authorization (provide a token number or an authorization number usually sent from the merchant via email or text).
- Virus, Spyware and Spam Prevention – Use of antiviruses and ensuring they are always updated.
- Avoid using bootlegs, keygens, and peer-to-peer sites
- Avoid visiting a link provided in an email, always type in the website.
- Bank and financial institution will never request for personal information over the mail.
- Avoid sharing information over public emails
- Avoid pop-ups and being wary of ads
- Always check the authenticity of a website using the WHOIS information.

7. CONCLUSION

While it is tedious and time consuming to always follow every guide and countermeasure before using a website. It however helps with protection. Based on the trends in the capabilities of phishing attacks, the following recommendations and countermeasures provide high-level non-technical guidance for ordinary users to help them deal with the increasing technical capabilities of criminals conducting phishing scams. Phishing is a highly profitable activity for criminals. And over the past years, as a result of increase in technology the different forms of phishing attacks are becoming more sophisticated. Even though awareness in response and countermeasure is also becoming

more accessible. It is still not anywhere in correlation to phishing attacks implementation. And every day we see innocent people being victims, all as a result of lack of awareness and basic knowledge of how to safe guard themselves from these criminals.

As a result, we attempted to enlighten the ordinary user about website forgery phishing attacks and also provide countermeasure. However website forgery based phishing attack is only a part of the various forms of phishing attack.

Further studies could explore the other types of phishing attacks such as man-in-the-middle, malware-based, key-loggers, web Trojans etc. and provide a non-technical guide and countermeasures that aim towards the ordinary user whom is not familiar with the IT know how.

REFERENCES

[1] Internet Users in the World Distribution by World Regions - 2014 Q2. Internet World Stats. N.p., 19 Mar. 2015. Retrieved from Web. 25 Apr. 2015.

[2] Avivah, L. *Phishing Attack Victims Likely Targets for Identity Theft, Gartner FirstTake*. FT-22-8873, Gartner Research, 4 May, 2004.

[3] Bleau, Heidi. RSA Monthly Online Fraud Report -- January 2014 (2014): n. pag. EMC.COM. EMC. Web.

[4] Jakobsson, Markus. "Modeling and preventing phishing attacks." *Financial Cryptography*. Vol. 5. 2005.

[5] Milletary, Jason, and CERT Coordination Center. "Technical trends in phishing attacks." *Retrieved December* 1.2007 (2005): 3-3.

[6] Kirda, Engin, and Christopher Kruegel. "Protecting users against phishing attacks with antiphish." *Computer Software and Applications Conference, 2005. COMPSAC 2005. 29th Annual International*. Vol. 1. IEEE, 2005.

[7] Chen, Juan, and Chuanxiong Guo. "Online detection and prevention of phishing attacks." *Communications and Networking in China, 2006. ChinaCom'06. First International Conference on*. IEEE, 2006.

[8] Prakash, Pawan, et al. "Phishnet: predictive blacklisting to detect phishing attacks." *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010.

[9] Moore, Tyler, and Richard Clayton. "An Empirical Analysis of the Current State of Phishing Attack and Defence." *WEIS*. 2007.

[10] Iliyev, D., and Yong Bin Sun. "Website forgery prevention." *Information Science and Applications (ICISA), 2010 International Conference on*. IEEE, 2010.

[11] Dhamija, Rachna, and J. Doug Tygar. "Phish and hips: Human interactive proofs to detect phishing attacks." *Human Interactive Proofs*. Springer Berlin Heidelberg, 2005. 127-141.

[12] Dhamija, Rachna, J. Doug Tygar, and Marti Hearst. "Why phishing works."*Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006.

[13] 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*. ACM, 2006.

[14] How Did I Get Infected? - MalwareTeks. How Did I Get Infected? - MalwareTeks. N.p., n.d. Web. 25 Apr. 2015. <http://www.malwareteks.com/articles/MalwareSources.php>.