Advances In Multidisciplinary & Scientific Research
A Multidisciplinary & Interdisciplinary Journal

Vol. 1, No. 2, September, 2015

Phishing & Anti-phishing Techniques: An Exposition

Durosinmi' A.E & Adenekan, O.A
[1&2]Department of Computer Engineering
Moshood Abiola Polyetchnic, Abeokuta, Ogun State. Nigeria.
cunlexie@hotmail.com, adenekanolujide@yahoo.com

Adekusibe, K. G.
Department of Computer Engineering
Ogun State Institute of Technology
Igbesa, Ogun State. Nigeria.
kusibus@yahoo.com

Alasiri, O. A
Department of Computer Engineering,
Cifman Institute of Technology and Management
Yaba, Lagos State. Nigeria.
olaitanalashiri@yahoo.com

ABSTRACT

Phishing is a current social engineering attack that results in online identity theft. Although a simple attack, phishing has become a large problem for organizations that are doing online business. The number of phishing scams is continuously growing, and the cost of the resulting damage is increasing. Researchers, as well as the IT industry, have identified the urgent need for anti-phishing solutions and recently, some solutions to mitigate phishing attacks have been proposed. We analyze these attacks and identify that most of them exploit the fact that users are not sufficiently aware of the secure site identification mechanisms in browsers. Some of these anti-phishing solutions are browser-based and have been implemented as plug-ins. In fact, it appears that even web designers are often confused about the need to identify login forms securely; we show several sites that prompt users for passwords using unprotected forms. In this paper we presented phishing in detail (including attack process and classification of phishing attack) and reviewed some of the existing anti- phishing techniques along with their advantages and disadvantages and also make recommendations on how both can be mitigated.

Keywords: Phishing, fraud, detection, E-mail, Security

Aims Research Journal Reference Format:
Durosinmi, A.E , Adenekan, O.A , Adekusibe, K. G. & Alasiri, O. A (2015): Phishing & Anti-phishing Techniques: An Exposition. Advances in Multidisciplinary (AIMS) Research Journal. Vol 1, No. 2 Pp 61-70.

1. INTRODUCTION

The term phishing was first used in the Internet literature in 1996 by the hacker group who stole America Online (AOL) accounts' credentials. Phishing is originated from Phreaking, which is considered as the science of breaking into phone networks using social engineering. A phishing attack is also based on the concept of social engineering in which users are tricked into opening malicious attachments or embedded links in the e-mails(Sood and Enbody, 2014). Phishing is the combination of social engineering and technical exploits designed to convince a victim to provide personal information, usually for monetary gains to the attacker (Phisher). Phishing scams can happen when malicious organizations or people (also known as cyber criminals) present themselves as an entity users can trust, and then try to trick them into providing personal information.

Phishing scams normally occur via emails, websites, text messages and phone calls that can delude recipients to think that Christmas came early. As explained in Busylel.ie (2014) and the Investor Publications of the US Securities and Exchange Commission (2013), cybercriminals will often pose as your bank or financial institution, your employer, or any other entity that you normally trust with your information. Only when the email phishing process and characteristics are fully understood can effective measures be designed against phishing attacks. Successful phishing attacks are based on a form of copying or re-engineering, a website's design and layout to pass themselves off as a genuine (targeted) website. A malicious website is crafted which looks and feels like the original site, convincing unsuspecting users that they are giving personal information to a trusted organization [Zdziarski, Yang and Judge, n.d.].

Zdziarski et al. continued with their findings that users are frequently drawn to the sites by forged emails designed to look like legitimate correspondence and may even copy the body from real email, but when the user clicks a link to visit the website, they will be directed to the malicious site instead. The more convincing phishing attack appears, or rather, the more genuine a malicious website looks, the more success the attack will have in extracting personal information. Some phishing attacks go so far as to create faux websites for which there is no legitimate counterpart; e.g. a page prompting users for personal information the organization would not have otherwise asked for. According to Anti-Phishing Working Group (APWG), phishing activities have been increasing, and most phishing websites are hosted in the US. In 2012, an average of over 25,000 unique phishing email reports was reported to the APWG. Also, the number of unique phishing sites detected exceeded 45,000 per month [Wike, 2013].

1.1 Steps in Phishing Attack

According to Shujun Li and Roland Schmitz (2009), all phishing attacks fit into the same general information flow as stated in the following steps.

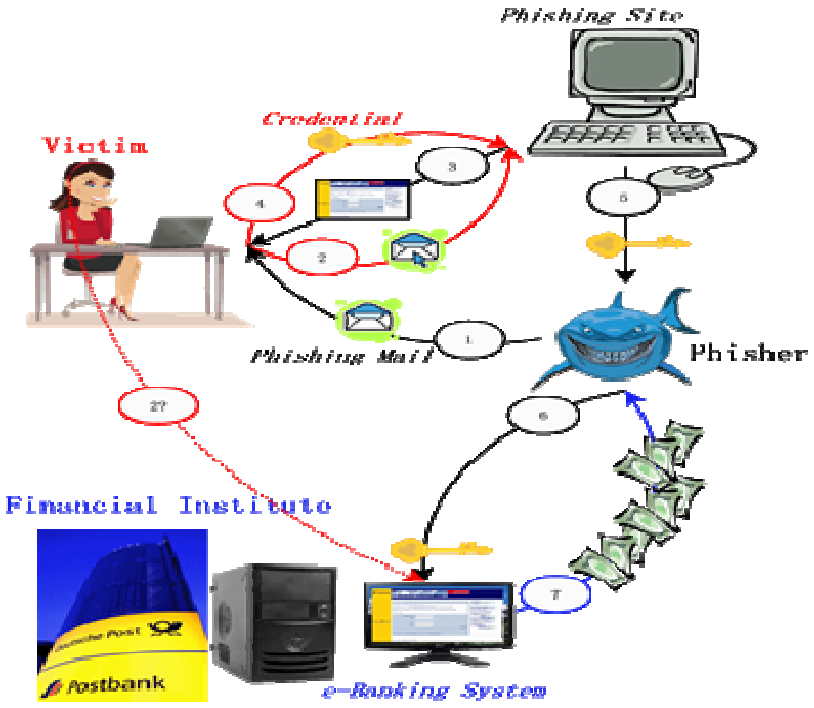

Fig. 1: Typical Phishing Scenario
Source: http://www.hooklee.com/default.asp?t=Honeypots4AntiPhishing

The steps are:

- **Step 1:** Blocking the information flow from phishers to potential victims phishing email detection and filtering, email authentication, anti-malware software, cousin domain rejection, and so on.
- **Step 2-4:** Avoiding credential leakage user education, phishing site warning, inconsistent DNS information detection, cross-site/injected script rejection, mutual authentication, trusted path between user and web browser, delayed password disclosure, and so on.
- **Step 5:** Preventing phishers from getting stolen credentials early detection and quick takedown of phishing sites, fake credential submission, password rescue, and so on.
- **Step 6:** making stolen credentials useless two-factor user authentication, password hashing, transaction monitoring and reconfirmation, and so on.
- **Step 7:** Preventing phishers from getting the stolen money, or catching phishers transaction authentication, intentional transaction delay, law enforcement, and so on.

2. PHISHING TYPES & TECHNIQUES

Many techniques and algorithms have been developed and implemented for the prevention of phishing and to secure the theft of confidential information (usernames, passwords, security key, credit card/debit card/master card details) [Tak and Ojha, 2013]. However, there are still many issues remaining on this matter. Many techniques and schemes, according to Gaurav Tak and Gaurav Ojha (2013), are being proposed to provide a secure environment for e-banking services, e-commerce services, and payment gateway services and to block the sniffing, eavesdropping, and so forth. So that transmission of the confidential information will be preserved, and unauthorized personnel cannot access that information. However, day by day, phishing attacks are increasing. While most phishing attacks target the financial transaction website (Banking site, e-commerce, e-shopping website, payment gateway websites), more and more phishing incidents targeting online game operators and large ISPs (internet service provider) have also been discovered.

2. 1 Fruit Sucker

This has to do with an attacker who breaks into a hacked, vulnerable website with an existing phishing page and changes the email address. Often since the attacker can see where the data is being emailed to he/she will keep the original email addresses intact and merely add his/her own email address in the BCC field. This is a very easy way for an attacker to make extra money. All passwords and account numbers keyed in reach his inbox directly without tipping off the original phishers. If the phishers are rookies and lack automated money transferring scripts, are too lazy to keep a watchful eye on their victims or are situated in a different time zone, then these advantages can help the fruit sucker withdraw a large amount of data (or other assets) from the compromised accounts before they do [Infosec Institute, 2012].

2.3 Spear Sucker

An attack against the original crackers. For example, a good guy who breaks into the phishing websites and changes the email address to the NEDBANK's CSO's or CEO's email address. After this, he contacts the bank to make them aware of the security breach [Infosec Institute, 2012].

2.4 Haxtortionist

When an attacker patches the system, pulls down the phishing page and emails the attackers threatening them that he/she will report the crime and inform NEDBANK of their malicious activity. Reporting such abuse to email servers hosted by Google, Yahoo and similarly large companies in the United States is easy. In this way, the attacker may extort a small share of money from the original crackers in return for keeping silent [Infosec Institute, 2012].

2.5 Robin-HAT

Here the attacker, after collecting many passwords, changes the recipient's email address for the purpose of redistributing wealth. He/She withdraws money from the accounts and donates a significant portion to charity. Such individuals cannot be called gray hats because they are criminals robbing from other criminals. They are Robin-HATS, those who steal from rich victims and their attackers and redistribute the wealth to the poor and needy [Infosec Institute, 2012].

2.6 Server-side exploits [Berghel, Carpinter and Jo, 2006]

Any discussion of the exploitation of server-side vulnerabilities to assist in a phishing attack quickly transcends phishing and enters the realm of general hacking and cracking; Suffice to say there are numerous techniques for exploiting operating systems, applications, and network protocols that a phisher could use if they were determined to comprise a legitimate website in order to conduct a phishing attack. However, two 'non-invasive' techniques of relevance to phishers will be discussed: cross-site scripting and preset sessions

a. Cross-site scripting (CSS or XSS) seeks to inject custom URLs or code into a web-based application data field and takes advantage of poorly developed systems. Three techniques are typically used. These are:

HTML substitution:

http://www.citibank.com/ebanking?URL=fakesite.com/login.htm

In this example, the standard legitimate website content is rendered, but the web application uses a parameter to identify where to load specific page content (for example the login box); in this case, that content is fetched from fakesite.com (whose URL could be obfuscated using previously described techniques).

Forced loading of external scripts:

http://www.citibank.com/ebanking?page=1&response=fakesite.com%21secretScript.js&go=2

In this example, a script to be executed is passed to the web application.

Inline embedding of active content:

http://www.citibank.com/ebanking?page=1&client=<SCRIPT>...
</SCRIPT>

In this example, the script is placed in the URL and executed by the web application.

Preset sessions use session identifiers. Session identifiers are typically used in HTTP and HTTPS transactions as a mechanism for tracking users through the website and for managing access to restricted resources (i.e. manage state). Session IDs can be implemented in a variety of ways; for example, cookies, hidden HTML fields or URL parameters. Most web applications allow the client to define the session ID. This allows the phisher to embed a session ID within the URL (that refers to the legitimate server) sent as part of the initial email. For example, https://mybank.com/ebanking?session=3V1L5e5510N. Once the email is sent, the phisher polls the legitimate server with the predefined session ID; once the user authenticates against the given session ID, the phisher will have access to all restricted content [Berghel, Carpinter and Jo, 2006].

2.7 Client-side vulnerabilities [Berghel, Carpinter and Jo, 2006]

Any discussion of client-side vulnerabilities is similar to that of its server-side counterpart: there is a multitude of vulnerabilities that a smart phisher could take advantage of to execute arbitrary code or to manipulate the browser. Given their exposure to the internet, it is not surprising browsers suffer from a significant number of security vulnerabilities. Most browsers also support some plug-ins, each of which carries its own security risks. While patches are typically available promptly, home users are notoriously poor at applying them quickly; therefore, phishers have ample time to exploit most security vulnerabilities if they choose to do so.

Some past exploits used by phishers include:

Microsoft Internet Explorer URL mishandling: a URL such as:
The real URL : http://www.citibank.com%01@fakesite.com/phisher.html
What the User sees: http://www.citibank.com
Where the browser goes: http://fakesite.com/fakepage.html

By inserting a %01 string in the username portion of the URL, the location bar displays http://www.citibank.com while redirecting the user to fakesite.com. Earthlink, Citibank, and PayPal were all targeted using this particular flaw.

Microsoft Internet Explorer and Windows Media Player combination
This vulnerability allowed the execution of a Java JAR archive, disguised as a Windows Media Player skin, which could access local files.

RealPlayer heap corruption:
RealPlayer is available as a plug-in for most browsers, and allows the user to view the proprietary RealMedia format. By creating a malformed RealMedia file, and embedding it in a website to ensure it is automatically played, it is possible to cause a heap corruption, which would allow the execution of arbitrary code.

While malware can often be eliminated with a regularly updated antivirus utility, browser (or any client-side) exploits cannot be defended against until a patch is available and applied.

3. ANTI-PHISHING TYPES & TECHNIQUES

Anti-phishing refers to the method employed in order to detect and prevent phishing attacks [Gaurav, Mishra and Jain, 2012]. Anti-phishing protects users from phishing Many techniques and algorithms have been developed and implemented for prevention of phishing and to secure the theft of confidential information (usernames, passwords, security key, credit card /debit card/master card details). However, there are still many issues remaining on this matter [Tak and Ojha 2013].

3.1 Phishing Mail Detection Techniques

Blacklists and heuristic are arguably the most popular phishing detection techniques. As evaluated in, although blacklists achieve low false positives, their detection rates suffer at zero-hours and are evaluated to detect only 20% of zero-hour phishing attacks. On the other hand, heuristics can detect constantly phishing attacks at a similar rate. However, heuristics was evaluated to have high false positives.

The effectiveness of a blacklist based solution depends on the time it takes until a phishing site is included. This is because many phishing pages are short-lived, and most of the damage is done in the time span. The techniques are described in detail below. The suspicious URL is matched against a list of known Phishing sites. This method is susceptible to “zero-day attacks”. Also, techniques like URL obfuscation and routing through alternate domain name can hinder this method ineffective [Hajgude and Ragha, 2013].

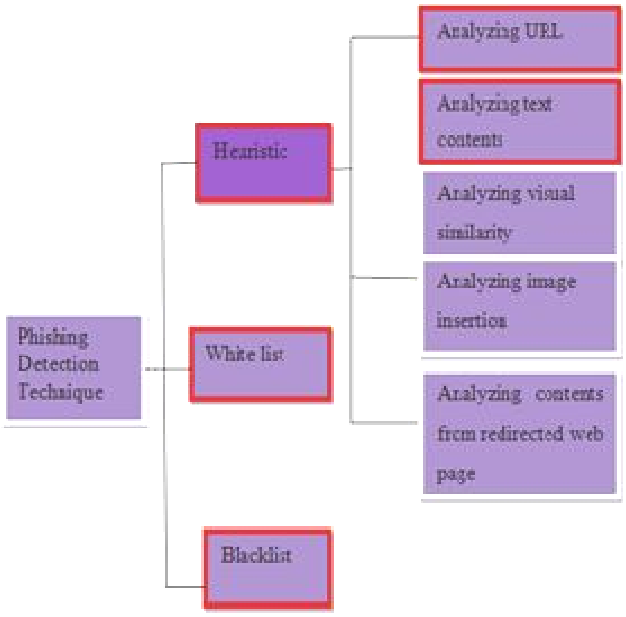

Fig. 2: Phishing Mail Detection Process

Most of the heuristics used are subjective and produce a large number of false positives. This solution is not limited to URL processing but also analyzes the page layout. Although some heuristics is used in this solution, they are used only in the pre-processing stages, and the actual phish detection is completely independent of them [Hajgude and Ragha, 2013].

3.1.1 Black Listing

Blacklist is a collection of known phishing Web sites/addresses published by trusted entities like google's and Microsoft's black list. It requires both a client & a server component. The client component is implemented as either an email or browser plug-in that interacts with a server component, which, in this case, is a public Web site that provides a list of known phishing sites [Gaurav, Mishra and Jain, 2012].

3.1.2 Symptom-Based Prevention-

Symptom-based prevention analyzes the content of each Web page the user visits and generates phishing alerts according to the type and number of symptoms detected [Gaurav, Mishra and Jain, 2012].

3.1.3 Domain Binding-

It is a client's browser based techniques where sensitive information (e.g., name, password) is bind to particular domains. It warns the user when he visits a domain to which user credential is not bind [Gaurav, Mishra & Jain, 2012] and [Marimuthu, Gopal, Mehta, Rajan and Boominathan, 2014].

3.2 An Identity-Based Anti-Phishing Techniques

This technique follows mutual authentication methodology where both user and online entity validates each other's identity during handshake. It is an anti- phishing technique that integrates partial credentials sharing and client filtering technique to prevent phishers from easily masquerading as legitimate online entities. As mutual authentication is followed, there would be no need for users to re-enter their credentials. Therefore, passwords are never exchanged between users and online entities except during the initial account setup process [Gaurav, Mishra and Jain, 2012].

Advantage: It provide mutual authentication for the server as well as client side. Using this techniques user does not to reveal his credential password in whole session except the first time when the session is initialized [Gaurav, Mishra and Jain, 2012].

Disadvantage: In identity-based anti-phishing if a hacker gains access to the client computer and disable the browser plug-in then the method will be compromised against phishing detection.

3.3 Character Based Anti-Phishing Approach

Many time phishers try to steal information of users by convincing them to click on the hyperlink that they embed into phishing email. A hyperlink has a structure as follows. <ahref="URI"> Anchor text <\a>
where 'URI' (universal resource identifiers) provides the actual link where the user will be directed and 'Anchor text' is the text that will be displayed in user's Web browser and represents the visual link.
The character based antiphishing technique uses characteristics of the hyperlink to detect phishing links [Nelson and Kavipriya, 2014].

3.4 Phishing Detection Based on Behavior of User

It is a novel approach to detect phishing websites based on analysis of users' online behaviors – i.e., the websites users have visited, and the data users have submitted to those websites. Such user behaviors cannot be manipulated freely by attackers; detection based on those data can not only achieve high accuracy [Dong, Clark and Jacob, 2014].

Advantages:

It detects essential characteristics of a phishing attack, namely that phishing web pages request authentication credentials. The details of how users may be manipulated may change with future phishing attacks, but the requesting of such details remains constant. It analyzes the identity of websites using both IP addresses and domain names, it can detect pharming attacks, which are undetectable by many existing systems. Being freelance of the incoming information suggests that low value in maintenance, the system does not want updating once attackers vary their techniques, and then we've got way fewer evasion techniques to cope with [Dong, Clark and Jacob, 2014].

Disadvantage

It cannot handle all types of authentication credentials. It can handle static type authentication credentials such as username, password, security questions, and the nlikes, but dynamic authentication credentials shared between users and the legitimate websites cannot be handled by the current implementation (e.g. one-time passwords) [Dong, Clark and Jacob, 2014].

4. MITIGATING PHISHING

4.1 Educating Your People

Many organizations conduct seminars and workshops on ethical hacking and Internet security to educate their employees. This can be a quality step towards security awareness though many of their employees may not take it seriously and may not follow the instructions given at the workshop/seminar. Those kinds of employees can be a potential target of attackers/phishers. There are some methods of educating your employees that we can think about. Logical awareness has to be built. First, they are given instructions to check the English. To respond to that, the bad guys started writing professional English that is more than 95% identical to the original website. Thus, victims got exploited. Then phishers started to use the lock symbol, keeping in mind that, even if some clever employee/person knows about SSL, she/he can be trapped. Phishers have done this by forging the symbol. They did it by putting lock icons in the URL (favicon) on the web pages. Banks started putting the last four digits of credit card or other bank account detail; in response to that, attackers also started putting the first four digits of those numbers that are constants in the card detail provided by any bank. Thus, persons got exploited again.

Mitigations: Logical awareness has to be raised. Customers have to think on their own about whether something is legal and legitimate or a fake. When this awareness rises within them, there won't be any need for workshops or seminars for ethical hacking awareness [INFOSEC, 2013].

4.2 Phishing Scam Alert Add-ons/Extensions

Many organizations have built toolbars that use a ton of problem-discovering and -solving methods to determine whether a URL is fake or not. Even Microsoft also used this feature, built into Internet Explorer The concept is like this. If server visits any known fake/phishing URL, then that toolbar turns red. If that phishing or fake site is the one suspect site, then it turns yellow. Nowadays some websites use "extended validation." This is a new type of certificate that is sold to the website only after the credentials are checked very carefully and particularly. If a browser toolbar finds this type of website, then it turns green.
The first method has already been broken by researchers. It was presented in a research paper whose link is mentioned in the references. That is a very unconventional and unusual semi-technical method for breaking into the victim's mind. It uses a "picture-in-picture" method. Here the phisher displays a picture of the browser with a green toolbar so that that the user thinks it is safe to visit and thus she/he is exploited.

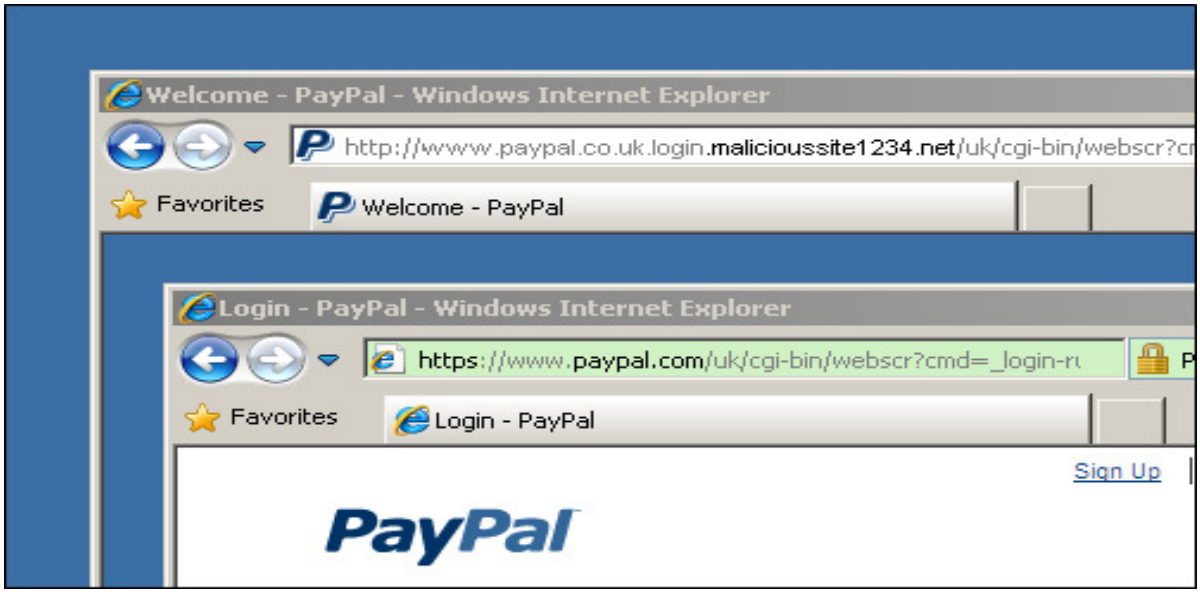

Fig. 3: Phishing Scam Alert Add-ons/Extensions

As can be clearly seen, the malicious URL is not https://www.paypal.com/uk, that is inside the browser's top window but it is displayed in the log-in window. The attacker also puts the favicon and outside logo to "prove" the legitimacy of his work. Thus, people think that this is the real page, and they log in to the website and their credentials are compromised. The second scenario, which is extended validation, can be broken by URL manipulation. Attackers use an almost identical URL, and they buy their own certificate and install it on their server. Now the URL of the phishing site and the original site are almost identical, as shown below:

Original Site:

www.chintanwov.com

Phishing Site: www.chintanvvov.com

As can be seen, in the first URL it's "wov" and in the second URL the attacker put "vvov"; "vv" looks like "w" and the client thinks that it's a genuine website and logs in. Thus, how their credentials get stolen and they get exploited. These types of phishing sites are called *"dodgy sites."* [INFOSEC, 2013]

4.2: Two-Factor Authentication

The two-factor authentication is also known as 2FA, two-step verification, or multi-factor authentication. It requires not only a username and password, but also some piece of information that only the user knows. That piece of information is known as a physical token. Using traditional credentials along with the physical token makes it very hard for a phisher to exploit his/her victim.

5. CONCLUDING REMARKS

In this paper, the phenomenon of phishing is growing, and the number of variations of techniques implemented demonstrates the high interest in these types of attacks by cybercrime. The phenomena must be carefully studied. Fundamental is training people in the secure use of computer tools, the cyber threat that is looming, and how the user can recognize threats to avoid serious problems [INFOSEC, 2012].

REFERENCES

1. Sood, A. and R. Enbody, (2014). "Targeted Cyber Attacks." SearchSecurity and Syngress. Retrieved from http://searchsecurity.techtarget.com/feature/Targeted-Cyber-Attacks
2. Busylel.ie, (2014). "Safety..." Retrieved from http://www.busylel.ie/Home/Page/safety
3. US Security and Exchange Commission, (2013). "'Phishing' Fraud: How to Avoid Getting Fried by Phony Phishermen." Investor Publications. Retrieved from http://www.sec.gov/investor/pubs/phishing.htm
4. Zdziarski, J., W. Yang and P. Judge, (n.d.). "Approaches to Phishing Identification Using Match And Probabilistic Digital Fingerprinting Techniques." Retrieved from http://www.mcafee.com/us/resources/white-papers/wp-approaches-to-phishing-identification.pdf
5. Wike, A. (2013). "Phishing Attacks And How To Prevent From Being Hooked." HongKiat. Retrieved from http://www.hongkiat.com/blog/phishing-reports-prevention/
6. Li, S. and R. Schmitz, (2009). **"**A Novel Anti-Phishing Framework Based on Honeypots.**"** Retrieved from http://www.hooklee.com/default.asp?t=Honeypots4AntiPhishing
7. Tak, G. and G Ojha, (2013). "Multi-Level Parsing Based Approach Against Phishing Attacks With The Help Of Knowledge Bases." International Journal of Network Security & Its Applications, Vol.5, No.6, November 2013, pp. 15-30. Retrieved from http://airccse.org/journal/nsa/5613nsa02.pdf
8. INFOSEC Institutes, (2012). "Attacking the Phishers: An Autopsy of Compromised Phishing Websites." Retrieved from http://resources.infosecinstitute.com/attacking-the-phishers/
9. Berghel, H., J. Carpinter and J. Jo, (2006). "Phish Phactors: Offensive and Defensive Strategies." Retrieved from http://www.berghel.net/publications/phishing/phishing.php
10. Gaurav, M. Mishra and A. Jain, (2012). "Anti-Phishing Techniques: A Review." International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 2, Issue 2, Mar-Apr 2012, pp. 350-355. Retrieved from http://www.ijera.com/papers/Vol2_issue2/BG22350355.pdf
11. Hajgude, J. and L. Ragha, (2013). "Performance Evaluation of Phish Mail Guard: Phishing Mail Detection Technique by using Textual and URL analysis." International Journal on Recent Trends in Engineering and Technology, Vol. 8, No. 1, Jan 2013, pp. 23-29. Retrieved from http://searchdl.org/public/journals/2013/IJRTET/8/1/61.pdf
12. Nelson, J. and R. Kavipriya, (2014). "Advanced Protection Against Phishing Pages By Using Multi-Usage Browser." International Journal of Scientific & Engineering Research, Volume 5, Issue 3, March-2014 431. Retrieved from http://www.ijser.org/paper/Advanced-Protection-Against-Phishing-Pages-By-Using-Multi-Usage-Browser.html

13. Marimuthu, K, D. Gopal, H. Mehta, A. Rajan, and P, Boominathan, (2014). "A Novel Way of Integrating Voice Recognition And One Time Passwords To Prevent Password Phishing Attacks." International Journal of Distributed and Parallel Systems (IJDPS) Vol.5, No.4, July 2014. Retrieved from http://airccse.org/journal/ijdps/papers/5414ijdps02.pdf
14. Dong, X, J. Clark and J. Jacob, (2014). "User Behaviour-Based Phishing Websites Detection." Proceedings of the International Multiconference on Computer Science and Information Technology, 2014, pp. 783–790. Retrieved from http://www.researchgate.net/publication/224370621_User_behaviour_based_phishing_websites_detection
15. INFOSEC Institute, (2013). "Phishing Counter-Measures Unleashed." Retrieved from http://resources.infosecinstitute.com/phishing-counter-measures-unleashed/
16. INFOSEC Institute, (2012). "Phishing: A Very Dangerous Cyber Threat." Retrieved from http://resources.infosecinstitute.com/phishing-dangerous-cyber-threat/