Phishing Identification and Prevention Techniques: Nimble Human Habits and Technology based Approach

(Paper under IT track)

Sagar More
Department of MCA, PK Technical Campus
MCA 2nd Year student (Semester 4).
Email : snmore1990@gmail.com
Mobile : 9970565243

Abstract

Phishing is a form of online identity theft that aims to steal sensitive information such as online banking passwords and credit card information from users. Phishing is a new type of network attack where the attacker creates a replica of an existing Web page to capture user's details. Phishing, an increasingly common form of online theft, is the act of tricking computer users into handing over control of their online accounts using, typically, a combination of a forged email and website. Phishing is a hacker's term that comes from the scams parallels with fishing, with the fake emails and website acting as the bait, and the victims accounts as the netted phish.

Phishing is done by spamming out authentic-looking emails that claim to come from a well-known financial or e-commerce institution such as Citibank, PayPal, eBay or America Online. These emails contain different messages, but usually follow the same formula: the recipient is asked to click on a link contained within the message, taking them to what appears to be a legitimate website. In fact, the website is a clever forgery, often virtually indistinguishable from the real thing.

According to a study by Gartner, 57 million US Internet users have identified the receipt of e-mail linked to phishing scams and about 2 million of them are estimated to have been tricked into giving away sensitive information. This paper presents a solution which can work as AntiPhish, that aims to protect users against spoofed web site based phishing attacks. To this end, AntiPhish techniques tracks the fake web pages and generates warnings whenever the user attempts to give away this information to a web site that is considered un-trusted.

1 : Introduction

The word 'Phishing' initially emerged in 1990s. The early hackers often use 'ph' to replace 'f' to produce new words in the hacker's community, Phishing is a new word produced from 'fishing', it refers to the act that the attacker allure users to visit a faked Web site by sending them faked e-mails, Phishing is an automated form of identity theft, targeted primarily at the casual e-mail user and stealthily get victim's personal information such as user name, password, and national security ID, etc. This information then can be used for future target advertisements or even identity theft attacks (e.g., transfer money from victims' bank account). In January of 2004, there were 174 phishing Web sites identified by the cross-vendor Anti-

Phishing Working Group. By December, there were over 1700. Finally, that year, the reported consumer loss due to Internet-based fraud was estimated between US$500 million (according to the Federal Trade Commission) and US$2 billion (according to the Anti-Phishing Working Group).

The frequently used attack method is to send e-mails to potential victims, which seemed to be sent by banks, online organizations, or ISPs. In these e-mails, they will make-up some causes, e.g. the password of your credit card had been mis-entered for many times, or they are providing upgrading services, to allure you visit their Web site to conform or modify your account number and password through the hyperlink provided in the e-mail. You will then be linked to a counterfeited Web site after clicking those links. The style, the functions performed, sometimes even the URL of these faked Web sites is similar to the real Web site. It's very difficult for you to know that you are actually visiting a malicious site. If you input the account number and password, the attackers then successfully collect the information at the server side, and is able to perform their next step actions with that information (e.g., send emails from your account, see your personal data & information, do cyber crime by using your identity, withdraw money out from your account).

Phishing attacks use a combination of social engineering and technical spoofing techniques to persuade users into giving away sensitive information (e.g., using a web form on a spoofed web page) that the attacker can then use to make a financial profit. Phishing scams have been receiving extensive press coverage because such attacks have been escalating in number and sophistication.

In general, phishing attacks are performed with the following four steps:

1) A fake web site which looks exactly like the legitimate Web site is set up by phisher
2) Phisher then send link to the fake web site in large amount of spoofed e-mails to target users in the name of legitimate companies and organizations, trying to convince the potential victims to visit their web sites.
3) Victims visit the fake web site by clicking on the link and input its useful information there.
4) Phishers then steal the personal information and perform their fraud such as transferring money from the victims' account.

This paper will detail the technical aspects of typical phishing campaigns, focusing on the tactics, methodology, and unique features of the phishing e-mail and the phishing Web site. It will outline how automated, inline network-based solutions, built on existing intrusion prevention technology, can be leveraged to assist network defenders protect their constituencies from online fraud.

2. Types of phishing attacks

In this section, we give a brief overview of the different types of phishing attacks to familiarize the reader with the threat. A real-world phishing attack is presented in Section 2.3.

2.1 Spoofing e-mails and web sites

Phishing attacks fall into several categories. The earliest form of phishing attacks were e-mail-based and they date back to the mid 90's. These attacks involved spoofed e-mails1 that was sent to users where attackers tried to persuade the victims to send back their passwords and account information. Although such attacks may be successful today, the success rate from the point of view of the attackers is lower because many users have learned not to send sensitive information via e-mail. A possible reason is that many security-sensitive organizations such as banks do not provide interactive services based on e-mail where the user has to provide a password. Most organizations, obviously, use their web sites for providing interactive services because they can rely on encryption technologies such as SSL. As a result, a typical user would find a request to send sensitive information such as a password via e-mail suspicious (especially considering the fact that many Internet users today receive a large number of spam e-mails from people that they do not know). Hence, many phishing attacks now rely on a more sophisticated combination of spoofed e-mails and web sites to steal information from victims. Such attacks are the most common form of phishing attacks today. In a typical attack, the attackers send a large number of spoofed e-mails that appear to be coming from a legitimate organization such as a bank to random users and urge them to update their personal information. The victims are then directed to a web site that is under the control of the attacker. This site looks and feels like the familiar online banking web site and users are asked to enter their personal information. Because the victims are directly interacting with a web site that they believe they know, the success rates of such attacks are much higher than e-mail- only phishing attempts. Besides e-mail, as an alternative form of message delivery, attackers have also started to use instant messaging systems such as ICQ or infrastructures such as Internet Relay Chat (IRC) to try to persuade and direct users to spoofed web sites. Once the victim follows a spoofed link, in order not to raise suspicion and to present the phishing web site as authentic as possible, attackers are employing various techniques. One example is the use of URLs and host names that are obfuscated and modeled so that they look legitimate to in- experienced users. Another example is the use of real logos and corporate identity elements from the legitimate web site. Some attacks also make use of hidden frames and images as well as Javascript code to control the way the page is rendered by the victim's browser.

2.2 Exploit-based phishing attacks

Some phishing attacks are technically more sophisticated and make use of well-known vulnerabilities in popular web browsers such the Internet Explorer to install malicious software (i.e., malware) that collects sensitive information about the victim. A key logger, for example, might be installed that logs all pressed keys whenever a user visits a certain online banking web site. Another possibility for the attacker could be to change the proxy settings of the user's browser so that all web traffic that the user initiates passes through the attacker's server (to perform a typical man-in-the-middle attack). Exploit-based phishing attacks are not the focus of our work in this paper. To mitigate exploit based phishing attacks (as well as other security threats that are directly related to browser security such as worms, trojans and spyware), browser manufacturers need to make sure that their software is bug free and that users are up to date on the latest security fixes. The focus of AntiPhish is to mitigate web site-based phishing attacks that aim to con victims into giving away their sensitive information.

2.3 : A real-world spoofed web site-based phishing attack example

On February 18th 2005, a mass e-mail was sent to thousands of Internet users asking them to verify their Huntington online banking account details. The e-mail claims that the bank has a new security system and that account verification is necessary. The attackers have supposedly inserted a legitimate URL https://onlinebanking.huntington.com/login.asp to the bank's online banking web site. However, the link actually points to a spoofed page on the server with the IP address 210.95.56.101. The aim of the attack is to steal the victim's account credentials, credit card information, and personal information such as the social security number. Once the victim enters the requested information, the phishing site redirects to the legitimate bank's web site.

The next section presents our browser extension (i.e., plug-in) for mitigating such phishing attacks.

3: Process of Phishing

According to Fig.1, Phishing starts with the external Email or web page having malicious intention of its owner. After getting phishing Email Novice user reads it and cannot be able to determine whether it is a fake email or original email. That Email contents prompts user to update or verify details of the account. These details can be username, password, name, address, PAN number, etc. After clicking the link given in Phished Email, user gets redirected to another webpage which is exactly similar to original webpage. If user enters the sensitive information into that fake page then all the entered details are immediately captured at server side of attacker, or attacker will get email with entered details.

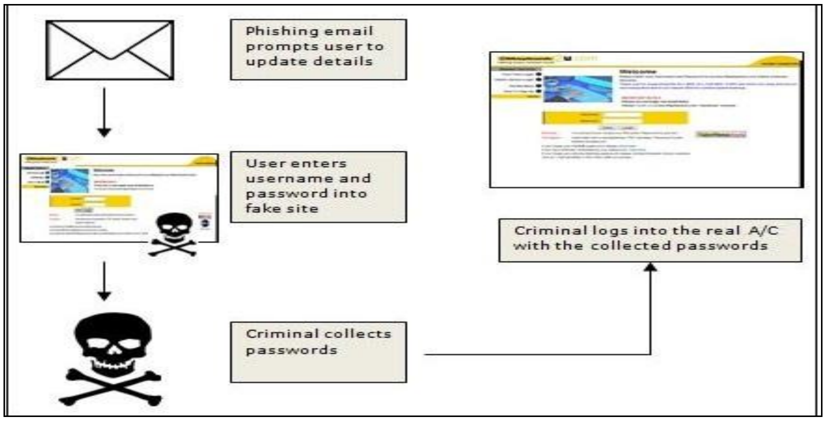
Phishing email prompts user to update details
User enters username and password into fake site
Criminal logs into the real A/C with the collected passwords
Criminal collects passwords

Fig. 1 : Process of Phishing

From the above figure it can be seen that .com sites are most vulnerable to phishing attacks. The figure also depicts that .net domain sites are also largely used by phisher for attack so it can be concluded that commercial site users becomes large victim of phishing attacks Further paper describes the literature survey and types of phishing and some of the anti-phishing techniques with their advantages and disadvantages

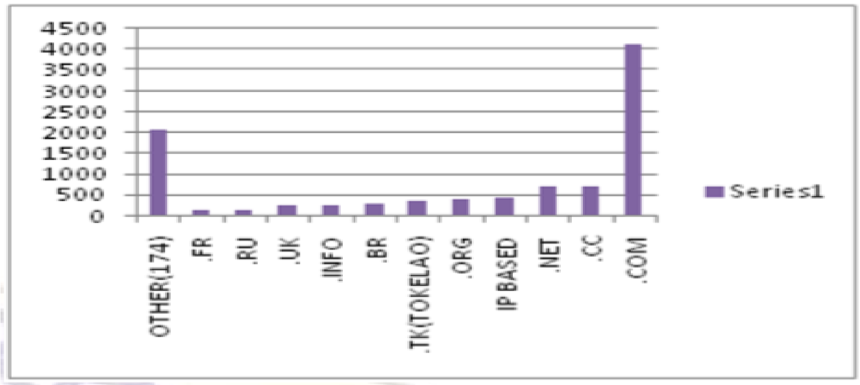

Fig 2: Domain name wise affect of Phishing

As shown in above given figure, here we can conclude that Financial sector of business is most targeted sector by attackers as compared to other sectors, Financial sector has 42.4% ratio of attacks. Retail/Services is second most targeted sector of attacks, attacks done on this sector has 20.5% ratio. Likewise ratios are Payment Services sector has 17.8%, Auctions has 4.1%, Gaming sector has 3.7%, Social Networking sector has 3.3%, ISP sector has 2.1%, Government sector has 1.2%, Classifieds sector has 1.1%, and commonly all other sectors has 3.8%. So, basically Finance sector has more risk of attacks, and Finance sector is most important element as it can affect nations economy.

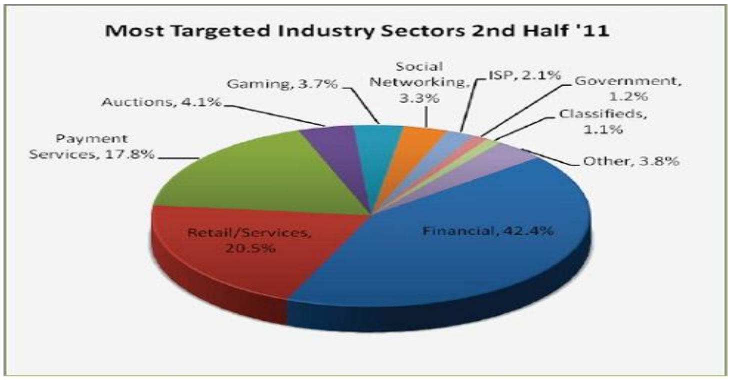

Fig 3: Phishing attacks ratio on different sectors

3. PHISHING ATTACKS CHARACTERISTICS

A. World Wide Web Consortium (W3C) objects:

A structured webpage is composed of W3C objects, some of these objects are [4]:

1) **Request URL (RURL):** External objects (such as images, external scripts, CSS) in a webpage are loaded from other URLs. For a normal corporate website, a large percent of those URLs are in its own domain. For instance,
<img src="https://home.peoplepc.com/i/60/common/ppco logo.gif"> in http://www .peoplepc.com.

2) **URL of Anchor (AURL):** A high portion of anchors in a legitimate webpage point to the same domain as the page itself. One example is <a href="http://www.ebay.com/"> in http://www.ebay.com. Webpage names must be meaningful to the visitor .Webpage names must be short, in all lower case, contain no spaces, use hyphens, not underscores between words.

3) **Server Form Handler (SFH):** For security and management reasons, most finance/e-business web portals require usernames and passwords. Therefore, those pages usually contain a server form handler. For example, <form action="/inetSearch/index.jsp" method="post" target=" top"> in http://www.chase.com. For phishing websites, the SFH usually is void or refers to a different domain.

B. Common properties of Phishing attacks:

The following lines represent number of properties of phishing attacks in the websites, they are:

1**)** **Logos:** The Phishing website uses logos found on the legitimate website to mimic its appearance. So phishers can load it from the legitimate website domain to their phishing websites (external domain).

2) **Suspicious URLs:** Phishing websites are located on servers that have no relation with the legitimate website. The phishing website"s URL may contain the legitimate website"s URL as a substring (http: //www.ebaymode.com), or may be similar to the legitimate URL (http://www.paypa1.com) in which the letter „L" in PayPal is substituted with number „1". IP addresses are sometimes used to mask the host name (http://25255255255/top.htm). Others use @ marks to make host names difficult to understand (http://ebay.com:top@255255255255/top.html) or contain suspicious usernames in their URLs (http://middleman/http://www.ebay.com).

3) **User input:** Phishing websites typically contain pages for the user to enter sensitive information, such as account number, password and so on.

4) **Short lived:** Most phishing websites are available for only a few hours or days – just enough time for the attacker to defraud a high enough number of users.

5) **Copies:** Attackers copy HTML from the legitimate websites and make minimal changes.

6) **Sloppiness or lack of familiarity with English:** Many Phishing pages have misspellings, grammatical errors, and inconsistencies.

4: Background

Ammar Almomani, et. al. [1] defines zero-day attacks as attacks that phisher mount using hosts that do not appear in blacklists and not trained on the old data sample and it is a noise data, which increases the level of difficulty to detect phishing email. They proposed a novel framework called phishing dynamic evolving neural fuzzy framework (PDENF), which adapts the evolving connectionist system (ECoS) based on a hybrid (supervised/unsupervised) learning approach. PDENF adaptive online is enhanced by offline learning to detect dynamically the phishing email included unknown zero-day phishing e-mails before it get to user account. PDENF is suggested to work for high-speed “life-long” learning with low memory footprint and minimizes the complexity of the rule base and configuration with few number of rules creation for email classification.

Madhuresh Mishra, et. al. [2] states various approaches to protect users against phishing, they proposed various anti-phishing techniques, that have been proposed that follows different strategies like client side and server side protection. In that paper they have studied 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.

Mona Ghotaish Alkhozae, et. al. [3] states Phishing is a type of Internet scams that seeks to get a use’s credentials by fraud websites, such as passwords, credit card numbers, bank account details and other sensitive information. There are some characteristics in webpage source code that distinguish phishing websites from legitimate websites and violate the w3c standards, so user can detect the phishing attacks by check the webpage and search for these characteristics in the source code file if it exists or not. They checked two webpage source codes for legitimate and phishing websites and compare the security percentages between them, they find the phishing website is less security percentage than the legitimate website; their approach can detect the phishing website based on checking phishing characteristics in the webpage source code.

5: PHISHING WEBSITES DETECTION METHODOLOGY

A. The phishing characteristics out of the W3C standards:

Phishers use some tricks and to fool users and tempting them, so our approach is to check for these tricks and factors in the webpage source code and calculate the security percentage based on these factors to classify the webpage if it is secure or not, they are:

1) **Https:** It is the secured protocol which used to tell us that this website is secured but it should be in "URL" of the website not in the body source of the webpage because phishers used Https inside their source code file to tell us that this images or this links is secured but it is not. The normal page should be like this <img src="mona.png" />. But there is some phishers use the SSL certificate in the source code like this <img src=https://www.xx.com/mona.png/>. They use https to make us think its secured website buts it is not. Many similar phishing attacks in which phishing websites use a certificate that can be expected to trigger a browser warning.

2) **Images:** All images in the website including website logo should load from the same URL of the website not from another website, so all links should be internal links not external links. Therefore, we check the links to detect any external links inside the source code like this : <img src=https: //www.Phishers.com/logo.jpg>" it is a phishing character.

3) **Suspicious URLs:** Most of the phishers use an IP address instead of using the actual domain name. Others use @ marks to ambiguous their host names.

4) **Domain:** It is the external domains mean: if we logged to website which its name is www.paypal.com and we found there is some URLS of links in the source code like this " www.pay-pal.com" which it is not the source URL so it means that this website try to hack our information .Phisher use forward domain also called domain redirection, it is a technique on the World Wide Web for making a webpage available under many URLs.

5) **Email:** There is a function on PHP called mail or email and it take our information which we enter in the forms like "MasterCard number, etc. "and send them when we press the pay button throw e-mail to the phishers e-mail. Phisher can insert PHP code inside Html code and use this function to send our information.

6) iframe: It is HTML tag code and used to embedding another webpage into current webpage. It creates a frame or window on a webpage so that another page can load inside this frame. Phishers use the iframe and make it invisible i.e. without frame borders, when the user goes to website, he/she cannot know that there is another page is also loading in the iframe window. It is a big problem which all people do not know it, it is like small website open in current webpage for example: we can open www.google.com in my page www.sagarmore.in by using iframe so when the people enter our website they will see the secured website is opened but it is not in the page it open throw iframe . Example: http://www.phisher.com/index.php?search='"><iframe src=http://google.com ></iframe> // Replace http://google.com by the phishing page.

7) Script: It is PHP files and there is some of phisher used scripts to send personal information or PC information to them, and some scripts send viruses or load from external websites. Scripts tag use to put any external file in the page like jquery or CSS and if it is with start and end tag, it is legal because this is the correct and standard script tag. Example: <script type="text/javascript" src="includes/jscripts/jquery.min.js"></script>, it is now load file to make the page appearance good. When there are tags like this <script> and between them any codes (not links) it is suspicious tags because this script code is javascript or any other languages which may be used to send personal information or PC information to phishers. So if we find <script> tags and there end tags </script> it is a legal tags, otherwise it is a phishing character.

8) Popup window: Phishers use popup windows to gain personal information. Often, these popups may ask to update, validate or confirm account information and it is likes official organizations websites. If the user enters his information in the popup windows, the phisher then steals this private information. There is two ways of popup window ,one of them is used to confirm or to tell us something and this window have a special way to write it in html or JavaScript like this < … onClick="window.open(sagar.html')"> and it is by html and legal window . The other way is illegal popup window because it is a javaScript file used in like this: "Open Popup" onClick="javascript:popUp(sagar.html')"> . It is illegal because it is open a new page from another website as registration page or ask the user information. It is a new full page and it open automatic when the user open a page or click on any link so it is illegal popup window.

6: Preventive Tips for Phishing Attack :

Here are five simple tips that will help Protect Your Online Accounts

a. Use strong passwords: Using strong passwords is the best way to ensure your social media accounts don't get hacked by a spammer or someone who wants to embarrass you. That's why it's so important to have a unique password for each of your accounts. Creating strong passwords is fairly easy – just remember to combine letters, numbers and symbols that require pressing the Shift key. It's also a good idea to change your passwords once in a while, like every two months.

Using strong passwords to **protect your online accounts** is a good start, but it will not protect you from serious hackers that rely on malware like keyloggers to steal your data. If your computer is infected with a keylogger, the hacker will have access to everything you type, including your passwords. Protecting your computer will help. You can also use strong passwords when protecting your home wireless connection.

Include post letters and numbers interchanged Try not to have "dictionary words." Basically try to have a random set of letters, like cga57g. If your password is made up of all words you can search for in the dictionary and find, someone can get your password using a dictionary hack.

- Make sure your password is long, about 15 characters in length. Otherwise, a hacker could use a brute force hack to get into your account.
- Spaces help, a lot. Even if your password is "theacp", making it "the acp", will increase the strength of it ten-fold.
- Don't use the same password for multiple accounts. Put a different number or symbol for each account added onto a base password if you can't remember different passwords. For example, "theacp" could turn into: theacp1,the@acp, the*acp, theacp!, etc…….
- Never enter your password into a page someone links you to, it may be a phishing site. Instead, make sure you go to the correct web page to be certain that your account information is secure.

b. Protect your computer

Protecting your computer is a vital part of securing your online accounts. You need to have solid security software to prevent more advanced hackers from accessing your online accounts and other sensitive data. There are many security suites available today, both paid and free. As a rule, it’s best to have a couple of different security programs installed on your computer, for example an anti-virus and an antimalware application. In addition to that, it’s always good to replace the Windows firewall with a more advanced one. Remember that you shouldn’t have two anti-virus programs installed at once because they might conflict with each other. Update your security software daily and scan your computer weekly to make sure there are no infections. Protect your online accounts include protecting your computer as they both rely on each other for safety.

c. Keep an eye on running processes

Every single program, whether visible or hidden, launches a process that is displayed in the Windows Task Manager. So if you think that your computer is infected, you should check running processes. It’s also good to monitor running processes on a regular basis. Windows 7 has a pretty decent task manager (press Ctrl+Shift+Esc to open it), but it’s still best to use a third party application, such as the free Sysinternal’s Process Explorer or Auslogics Task Manager. These programs provide more details than the built-in task manager and can help you nip malware in the bud. Some infections have a habit of masking themselves as Windows processes, such as svchost.exe and lsass.exe. So it’s always good to check your processes on Fileinspet.com, a Windows process library, and check their path.

c. Download with care : Do you like downloading free stuff from the Internet? I bet you do. But sometimes downloading free stuff can be dangerous. While a lot of free downloads are perfectly fine and come from legitimate sources, many are infected. They are designed to wreak your computer and steal your data. Never download anything that looks suspicious and stick to legitimate free downloads, be it software, songs, or videos.

d. Be careful when using unprotected public networks

We all love using free Wi-Fi in cafes and libraries. While they are great for browsing the web and reading the news, it’s not a good idea to use them for online banking, shopping and sometimes even email. These networks are unprotected, which means that a hacker sitting in the same cafe can easily access all of your open accounts and steal your passwords using special software. That’s why you should always take extra care and watch out for any strange activity. GMail users are lucky, because GMail tells you if more than two computers are using your account at the same time – just look below your messages to access this information.

e. Do not use Shortcut methods

There are some famous nimble habits of humans that is humans use shortcut methods to achieve results within less time, but these shortcut methods can affect to the owner of online account. Peoples have habit of clicking on link instead of typing it. A link can contain malicious Phishing link, so if user is clicking on that link then visual address of that link may correct but internally in networking the address behind the link can have different address that may not be known to user at the time of clicking on it.

Example : <a href=”www.fakewebsite.com”>www.realwebsite.com</a>

In above example user can only see www.realwebsite.com that is a original address of the site where user want to log on, but behind this link the actual address given by attacker is www.fakewebsite.com which is a fake website name readable by computers. So in this situation user gets redirected to a fake website which is similar in visual part of real website, but if user enters any details such as username, password, address, bank account number etc then this information is captured by a hacker or phisher.

f. Manually enter URLs :

As described in above point that shortcut methods may harm users online identity, So here we will cover the advantage of manually entering we address. People shows laziness to enter a website address manually due to this they use some shortcuts, but if user enters the website address manually then risk of being hacked is reduced, because at that point you will not be redirected to unwanted phishing page which is exactly looking as original web page. So always enter URLs manually.

g. Install AntiPhish Software :

Your system must have latest AntiPhish software installed so that software can detect your system’s web browser activity for tracking phishing activity and if any malicious activity is found by that software then it prompts user with a warning message about phishing attack. In

Mozilla Firefox browser there are lots of Plug-ins are available to prevent phishing attacks user can install that plug-ins instead of heavy AntiPhish software.

7: Conclusion

In this paper we proposed a phishing detection approach that determine the webpage security by identifying the webpage URLs, Key-logger Software, Popup, scripts, Iframe, Images, Emails etc. Finally we get noticed about original and fake web pages and Emails, the genuine web pages indicates secure website and others indicates the website is most likely to be a phishing website. We can Install AntiPhish software to prevent phishing attacks; This paper gives emphasis on user knowledge about technologies that is user must know current technology trends and techniques to prevent ourselves from such Phishing attacks. Also, we can install a browser plug-in to check the web pages and informs the user if there any possible attack.

Acknowledgement

We are thankful to Dr. Manoj Devare, Head of MCA department at PK Technical campus, Chakan, Pune. Without his guidance, motivation and support we would not have submitted this paper.

References:

1. Ammar Almomani et. al., “Phishing Dynamic Evolving Neural Fuzzy Framework for Online Detection Zero-day Phishing Email”, Indian Journal of Science and Technology, Vol: 6 Issue: 1 January 2013 ISSN:0974-6846, pp. 3960-3964.
2. Madhuresh Mishra. et. al., “Anti-Phishing Techniques: A Review”, International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622, Vol. 2, Issue 2,Mar-Apr 2012, pp.350-355
3. Mona Ghotaish. et. al., “Phishing Websites Detection based on Phishing Characteristics in the Webpage Source Code”, International Journal of Information and Communication Technology Research ISSN-2223-4985, Volume 1 No. 6, October 2011, pp.283-291
4. Phishing activity trend report 1st half /2011, http://www.antiphishing.org. [accessed on 10/02/2013]