Volume 3, Issue 5, May 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Phishing & Anti-Phishing Techniques: Case Study

Jyoti Chhikara
CSE Dept, PDMCEW
India.

Ritu Dahiya
CSE Dept, PDMCEW
India.

Neha Garg
CSE Dept, PDMCEW
India.

Monika Rani
CSE Dept,PDMCEW
India.

Abstract*— Phishing is a con game that scammers use to collect personal information from unsuspecting users. The false e-mails often look surprisingly legitimate and even the Web pages where users are asked to enter their information may look real. Phishing is similar to fishing in a lake, but instead of trying to capture fish, phishers attempt to steal personal information. This paper gives brief information about phishing, its attacks, steps that users can take to safeguard their confidential information. This paper also shows a survey conducted by netcraft on phishing.

***Keywords*— anti-phishing technologies, identity theft, Network security, Phishing attacks.**

I. Introduction

Internet has changed the life of human significantly and it has dominated many fields including e-Commerce, e-Healthcare etc. Internet increases the comfort of human life; on the other hand it also increases the need for security measures too. For example all web browsers and servers take almost every care to make guarantee the safe business through internet. Still they are vulnerable to attacks such as phishing. 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 scams have been receiving extensive press coverage because such attacks have been escalating in number and sophistication. Phishing is not limited to the most common attack in which targets are sent spoofed (and often poorly spelt) messages imploring them to divulge private information. Instead and as recently documented both in academic and criminal aspects, phishing is a multi-faceted techno-social problem for which there is no known single silver bullet. As a result of these insights, an increasing number of researchers and practitioners are attempting to quantify risks and degrees of vulnerabilities in order to understand where to focus protective measures.

A. Classification of Phishing Attacks:

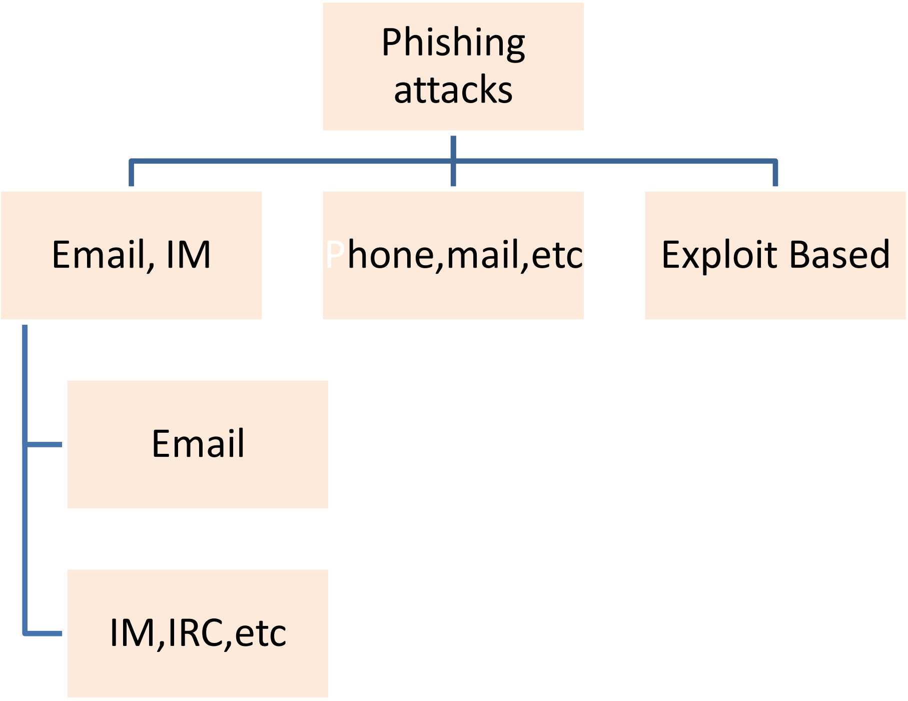


Fig1. Phishing attacks

- Spooled e-mails are sent to a set of victims asking them (usually) to upgrade their passwords, data accounts, etc.
- MSN, ICQ, AOL and other IM channels are used to reach the victims. Social engineering techniques are used to gain victim's sensitive information.
- Calling the victims on the phone, classic social engineering techniques are used by phishers.
- Another kind of attack is based on internet vulnerabilities. This approach is usually used to automatically install dialers.

B. Typical Process of Phishing:In a typical phishing attack[1], phishers send a large number of spooled emails to random no. of internet users that seem to be coming from a legitimate organization. Email urges to provide sensitive information. By clicking on the link provided in the mail, user is directed to a bogus site implemented by the attacker.

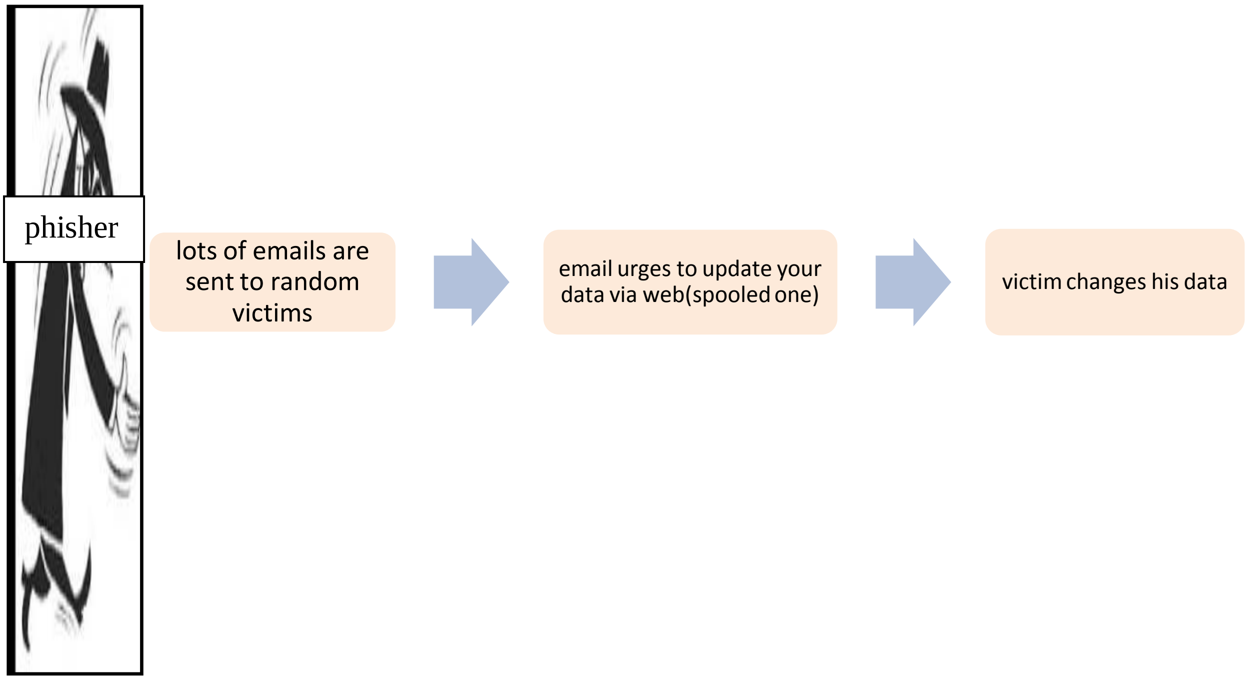

phisher
lots of emails are sent to random victims
email urges to update your data via web(spooled one)
victim changes his data

Fig2. Phishing process

C. Phishing Attack Stages

Phishing attacks involve several stages:

• The attacker obtains E-mail addresses for the intended victims. These could be guessed or obtained from a variety of sources.

• The attacker generates an E-mail that appears legitimate and requests the recipient to perform some action.

• The attacker sends the E-mail to the intended victims in a way that appears legitimate and obscures the true source.

• Depending on the content of the E-mail, the recipient opens a malicious attachment, completes a form, or visits a web site.

• The attacker harvests the victim's sensitive information and may exploit it in the future.

As shown in Figure 1 below, the phishing attack starts with an E-mail to the intended victims. The attacker creates the E-mail with the initial goal of getting the recipient to believe that the E-mail *might* be legitimate and should be opened. Attackers obtain E-mail addresses from a variety of sources, including semi-random generation, skimming them from Internet sources, and address lists that the user believed to be private [CNET]. Spam filtering can block many of the phishing Emails. If the institution whose customers are being phished regularly uses authenticated E-mail (such as PGP or S/MIME), the recipient may notice that the E-mail does not have a valid signature, thereby stopping the attack. Once the E-mail is opened by the user, the E-mail contents have to be sufficiently realistic to cause the recipient to follow the directions in the Email.

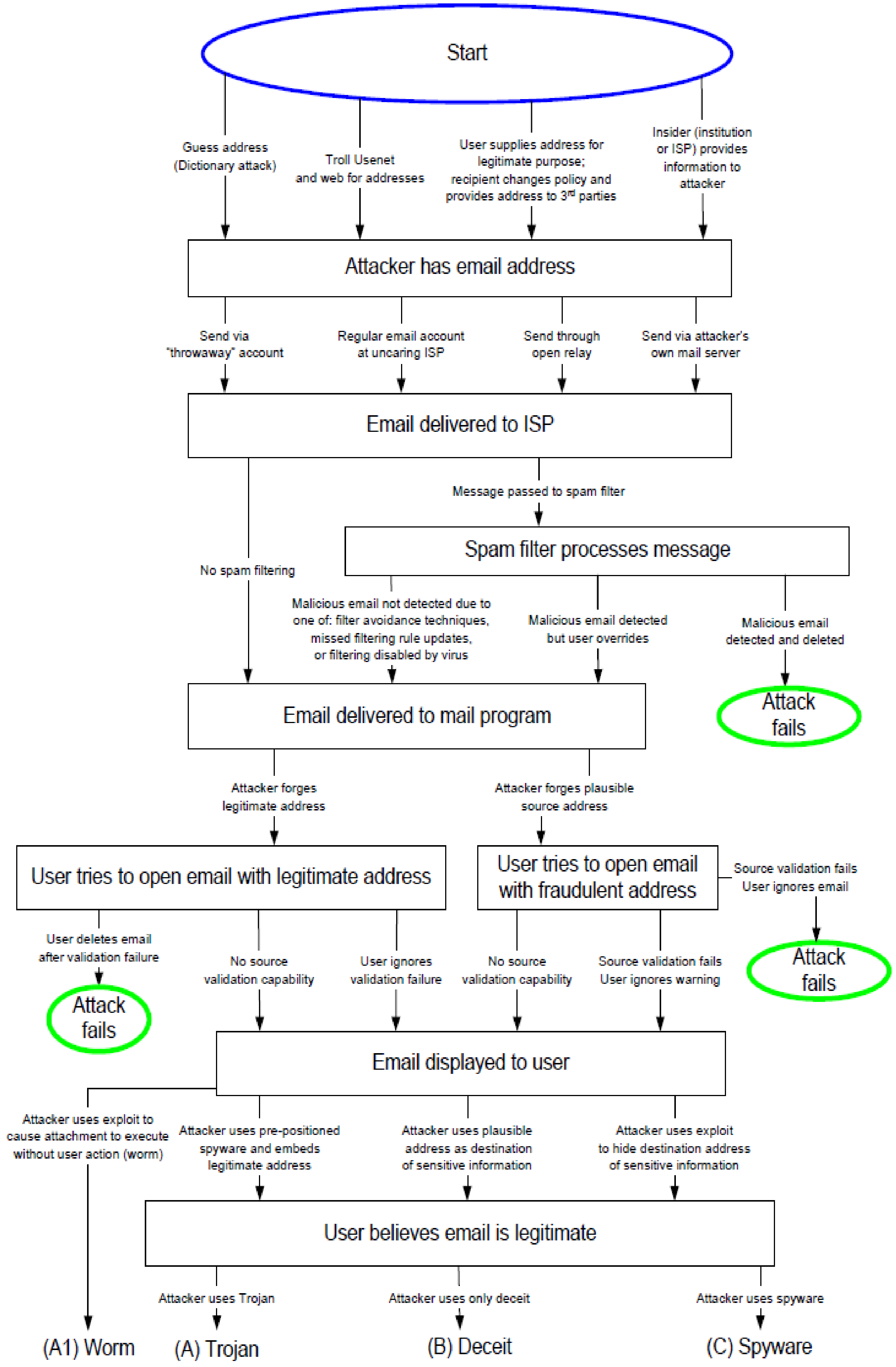

Start
Guess address (Dictionary attack)
Troll Usenet and web for addresses
User supplies address for legitimate purpose; recipient changes policy and provides address to 3rd parties
Insider (institution or ISP) provides information to attacker
Attacker has email address
Send via "throwaway" account
Regular email account at uncaring ISP
Send through open relay
Send via attacker's own mail server
Email delivered to ISP
Message passed to spam filter
No spam filtering
Spam filter processes message
Malicious email not detected due to one of: filter avoidance techniques, missed filtering rule updates, or filtering disabled by virus
Malicious email detected but user overrides
Malicious email detected and deleted
Attack fails
Email delivered to mail program
Attacker forges legitimate address
Attacker forges plausible source address
User tries to open email with legitimate address
User tries to open email with fraudulent address
Source validation fails User ignores email
Attack fails
User deletes email after validation failure
Attack fails
No source validation capability
User ignores validation failure
No source validation capability
Source validation fails User ignores warning
Email displayed to user
Attacker uses exploit to cause attachment to execute without user action (worm)
Attacker uses pre-positioned spyware and embeds legitimate address
Attacker uses plausible address as destination of sensitive information
Attacker uses exploit to hide destination address of sensitive information
User believes email is legitimate
Attacker uses Trojan
Attacker uses only deceit
Attacker uses spyware
(A1) Worm
(A) Trojan
(B) Deceit
(C) Spyware

Fig3. Stages in phishing attack

II. ANTIPHISHING TECHNIQUES

Antiphishing defenses can be server and client based solutions.

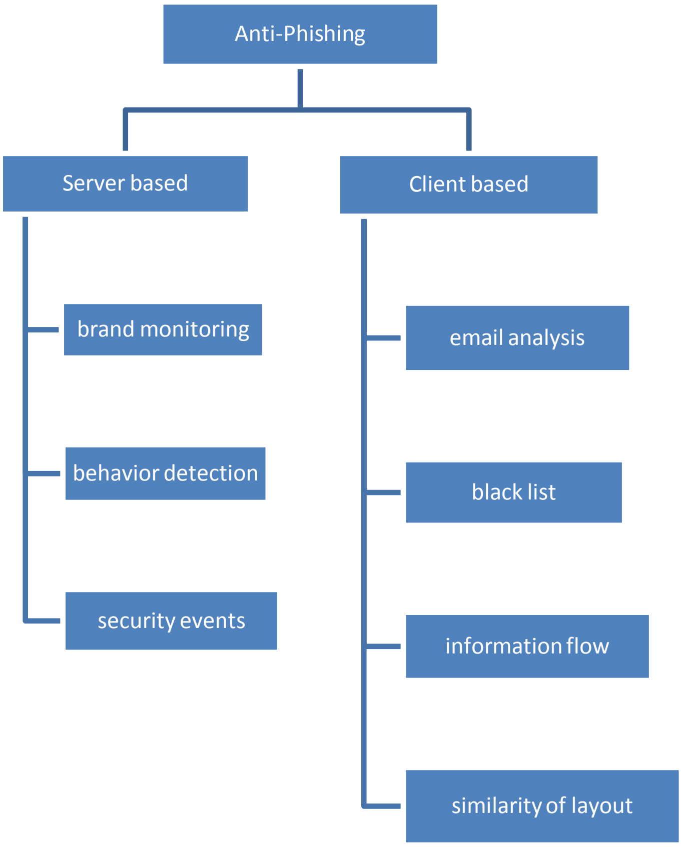


Fig4. Anti-phishing techniques

Server Based- these techniques are implemented by service providers (ISP, etc) and are of following types:

- Brand Monitoring: Cloning online websites to identify "clones" which are considered phishing pages. Suspected websites are added to centralized "black list".
- Behavior Detection: for each customer, a profile is identified (after a training period) which is used to detect anomalies in the behavior of users.
- Security Event Monitoring: security event analysis and correlation using registered events provided by several sources (OS, application, network device) to identify anomalous activity or for post mortem analysis following an attack or a fraud.

Client Based-these techniques are implemented on user's end point through browser plug-ins or email clients and are of following types:

- Email based analysis: email based approaches typically use filters and content analysis. If trained regularly, Bayesian filters are actually quite effective in intercepting both spamming and phishing e-mails. Bayesian algorithm explains the working of Bayesian filter:

Bayesian Algorithm:

1)Split e-mail in tokens.

- Need number of messages for spam and legitimate.
- Need frequency of each word for each type.

2)Calculate probabilities.

- P (legitimate) = word frequency /number of legitimate messages.
- P (spam) = word frequency/ number of spam messages.

3)Calculate likelihood of being spam (spamicity) using a special form of Bayes' Rule where likelihood = a/(a + b), where a is the probability of a legitimate word and b is the probability of spam word.

4) Choose tokens whose combine probability is farthest from 0.5 either way. This is because the farther it is from 0.5 (neutral), with more certainty we can say it belongs to either strategy.

- Do this for n numbers for n instance
- Combine their probability to get a figure for message using Bayes' Rule. In basic terms, Baye's Rule determines the probability of an event occurring based on the probabilities of two or more independent evidentiary events. For three evidentiary events a, b, and c, the probability is equal to

$$\frac{a\ b\ c}{}$$

abc+ (1-a)*(1-b)*(1-c)

- If the end result is closer to 1.0, then the message is classified as spam, and if it is closer to 0.0, the message is classified as legitimate.

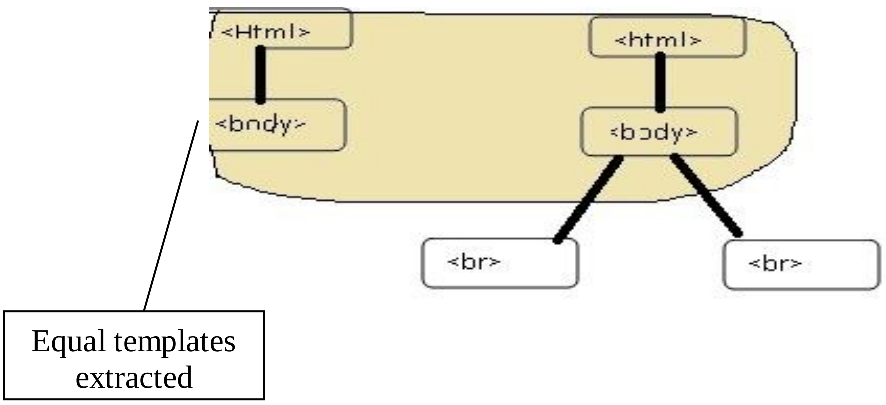
- Black Lists: black lists are collection of urls identified as malicious. The black-list is queried by the browser at run time whenever a page is loaded. If the currently visited url is included in the black list, the user is advised of the danger otherwise the page is considered legitimate.
- Information flow: information flow solutions are based on the premise that while a user can be easily fooled by URL obfuscation or a fake domain name, a program will not run. AntiPhish is an example of this type of technique which keeps track of sensitive information that the user enters into web forms, raising an alert if something is considered unsafe.
- Similarity of layouts: most advanced techniques try to distinguish a phishing page from a legitimate page by comparing their visual similarities. DOM-Antiphish computes the similarity value extracting the DOM-tree of the considered webpages.

  DOM-Antiphish description: when a password associated with certain domain is reused on another domain the system compares layout of current page with the page where the sensitive information was originally entered. For the comparison, DOM trees of the original webpage and the new one are checked. If the system determines that both trees are same in appearance, then phishing attack is assumed.

  DOM-AntiPhish Similarity Computation:

  -----phishing example---

Legitimate web page	Phishing web page
<html> <body><br>Hello<br></body> </html>	<html> <body> <br>Hello</br> </body> </html>
Legitimate DOM tree	Phishing DOM tree



III. PHISHING PROTECTION BEST PRACTICES

BOTNETS

In today's business and consumer computing paradigm, an emerging tool for various malicious activities is the botnet. Botnets—networks of compromised machines infected with malicious programs—have been identified as a leading cause for phishing, a serious form of spam.

Bots :-A bot—short for robot—is an automated software program that operates as an agent for a user or another program, or alternatively, simulates human activity. On the Internet, the most ubiquitous bots—more commonly known as spiders or Web crawlers—are legitimate programs that access Web sites and collect content for search engine databases. Bots have also been created to verify stock quotes or compare prices on shopping-based Web sites. Other bots such as knowbots and chatterbots have been used in a variety of legitimate ways.

However, bots are increasingly used for malicious purposes; these are known as IRC (Internet Relay Chat) bots. This type of bot is created when a computer virus or worm installs a backdoor program—such as a Trojan horse (a malicious program disguised as, or embedded within, legitimate software) or a drive-by downloader (which exploits Web browsers, e-mail clients, or operating system bugs to download malware without requiring any user intervention)—that leaves a PC Internet port open. The MyDoom (2004) and SoBig (2003) email worms, for example, employed this tactic. The infected machines subsequently become available for future activation. A hacker then searches for infected PCs with open ports.

Once located, the hacker installs the bot program onto their hard drives. The bot then typically connects to Internet Relay Chat to listen for commands, and the controller (a malicious third party) can unleash the effects of the bot by sending a single command to those machines. Bots can also be formed when their creators embed malware on Web pages; creators commonly use pornography, celebrity, Web hosting, or social networking Web sites for this purpose. Users unknowingly download the malware either by clicking on links containing the code or, worse, simply by visiting a URL.

Businesses and consumers can protect themselves from the devastating effects of phishing due to botnet activities in two ways: educating themselves about phishing techniques and employing technology solutions that combat phishing. The following checklist is a general best practice prescription for guarding against malicious threats:

Anti-Phishing Best Practice Checklist

Table 1. Anti-Phishing Best Practice Checklist

BEST PRACTICE	BUSINESS	CONSUMER
Always install, update, and maintain firewall and intrusion detection software including those that provide malware security.	✓	✓
Use latest web browser version and install security patches when available.	✓	✓
Practice awareness when receiving emails asking for account details	✓	✓
Never email financial/personal information.	✓	✓
Only open email attachments from trusted parties.	✓	✓
Never click on links in suspicious emails.	✓	✓
Report suspicious emails to appropriate authority.	✓	✓
Monitor logs from firewalls, intrusion detection systems, DNS servers, proxy servers on a daily basis for a signs of infections.	✓	
Monitor outbounds SMTP connection attempts that do not originate from normal SMTP mail gateways.	✓	
Establish rigorous password policies for clients, servers, routers and enforce them.	✓	
Ensure that approved devices can connect to the organization’s network.	✓	
Regularly read the latest news and info regarding phishing.	✓	✓

In terms of specific technologies, businesses and consumers alike should look for layered solutions that protect against both sending—that is, becoming an unwitting accomplice to propagating spam—and receiving phishing emails. From a business perspective especially, layered solutions should also offer content protection at the client side, or end points, and at the network gateway—as well as monitor network behavior. This ensures against “rogue” devices such as laptops and notebooks—which are not always under administrators’ control and may not have adequate or updated threat protection installed—infecting the entire network. The following checklist can serve as a guideline in making technology-related decisions to combat phishing:

Specific Anti-Phishing Technology Checklist

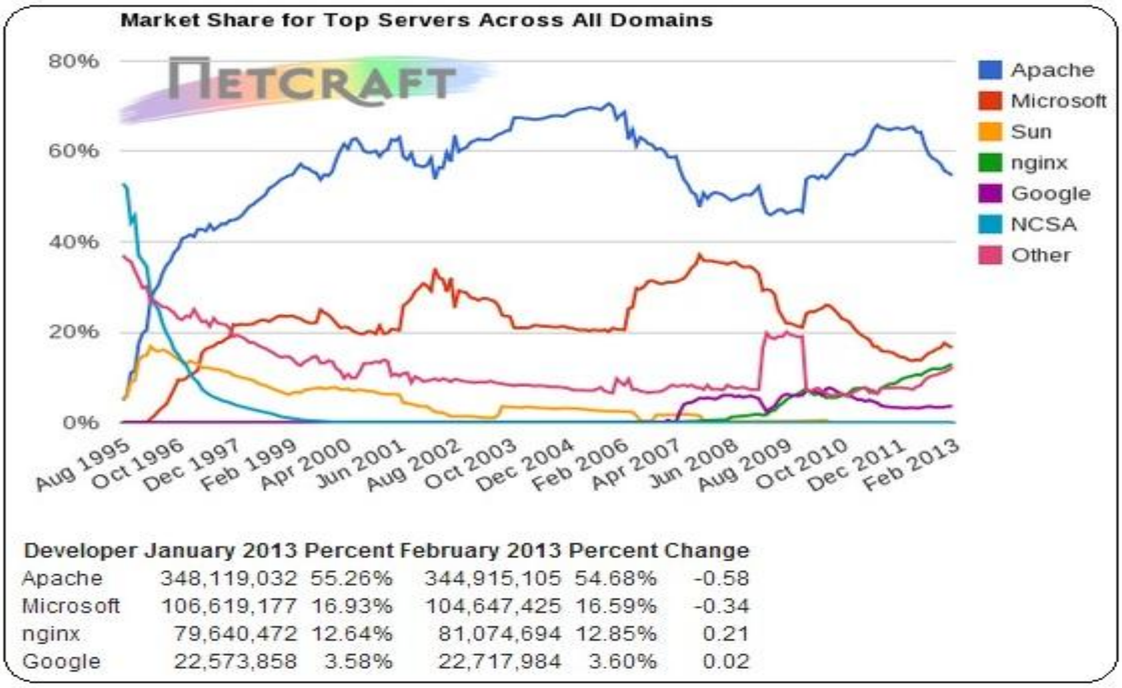
Table2. Specific Anti-Phishing Technology Checklist

Protection type	Protects against	
	Sending	receiving
Client side/end point	• Personal firewall • antivirus	• personal firewall • anti-virus • anti-phishing toolbar/enabled browser
Network behavior	• intrusion detection system(IDS) • intrusion protection system(IPS) • network content inspection	• IDS/IPS • network content inspection

Network gateway	• firewall • gateway anti-virus • gateway anti-spam	• domain reputation measurement

IV. NETCRAFT's WEB SERVER SURVEY 2013

Netcraft, an Internet services company that provides web hosting & web server analysis & has launched its February 2013 web server survey after responses from over 630,790,500 web sites.



Developer	January 2013	Percent	February 2013	Percent	Change
Apache	348,119,032	55.26%	344,915,105	54.68%	-0.58
Microsoft	106,619,177	16.93%	104,647,425	16.59%	-0.34
nginx	79,640,472	12.64%	81,074,694	12.85%	0.21
Google	22,573,858	3.58%	22,717,984	3.60%	0.02

Fig5. Netcraft's survey

There was a major decline this month of sites that use Microsoft IIS & Apache, with both servers seeing a combination or more than 5 million hostnames.

On the other side, nginx saw a 12.85% increase in business in last month January with 1.4 million more hostnames than December. The largest gains in hostnames positions nginx as one of the most well-known webservers, placing it less than 500 individual sites as Microsoft's IIS, which also has under 13 % of business. Tengine, an nginx derivative managed by China e-tailer Taobao, and now is used for just about 4 million hostnames. For the meantime, Alibaba, which is affiliated with Taobao, has the second largest number of hostnames in China, with more than 11% of the hostnames in China. Although China makes up 19% of world population, only 5.8% of the world's websites are actually hosted in China. Still leading China, Microsoft has 38% of Chinese hosted websites using IIS, followed by 26% using Apache, & lastly 1% that uses nginx, which is considerably above average. In a report, Netcraft stated that taobao draws the second highest number if phishing attacks next to Facebook.

Netcraft is reporting blocking nearly 6,000 urls targeting taobao users.

V. REPORT PHISHING

Whenever user finds a particular webpage as spam one he can report the phishing on following websites:

Businesses and consumers can file phishing reports with the following organizations:

Anti-Phishing Working Group
http://www.antiphishing.org
Digital Phishnet
http://www.digitalphishnet.org/
Federal Trade Commission
http://www.consumer.gov/idtheft/
Internet Crime Complaint Center (a joint project of the FBI and the National Collar Crime Center)
http://www.ic3.gov
Trend Micro Anti-Fraud Unit
antifraud@support.trendmicro.com

VI. CONCLUSIONS

Phishing differs from traditional scams primarily in the scale of the fraud that can be committed.

In order to combat phishing, business and consumers need to adopt best practices and practice awareness, educate themselves about phishing and anti-phishing techniques, use current security protection and protocols, and report suspicious activities. By doing so, they can reduce their exposure to fraud and identity theft, safeguard their confidential information, and help fight one of today's most serious and ongoing threats of phishing. The most effective solution to phishing is training users not to blindly follow links to web sites where they have to enter sensitive information such as passwords. The final technical solution to phishing involves significant infrastructure changes in the Internet that are beyond the ability of any one institution to deploy. However, there are steps that can be taken now to reduce the consumer's vulnerability to phishing attacks. Some of those steps are:

For Corporations:

- Establish corporate policies and communicate them to consumers.
- Provide a way for the consumer to validate that the E-mail is legitimate.
- Stronger authentication at web sites.
- Monitor the Internet for potential phishing web sites.
- Implement good quality anti-virus, content filtering and anti-spam solutions at the Internet gateway.

For Consumers:

- Automatically block malicious/fraudulent E-mail.
- Automatically detect and delete malicious software.
- Automatically block outgoing delivery of sensitive information to malicious parties.
- Be suspicious.

All of these technologies are available now and can be deployed by both consumers and institutions interested in protecting their customers.

REFERENCES

[1] Angelo P.E. Rosiello, Engin Kirda Christopher kruegel and Fabrizio Ferrandi."*A Layout Similarity Based Approach For Detecting Phishing Pages*". IEEE Conference on Security and Privacy in Communication Networks, Nice, France, September 2007.

[2] Neil Chou, Robert Ledesma, Yuka Teraguchi, Dan Boneh, and John Mitchell."*Client-side defense against web-based identity theft*".In 11th Annual Network and Distributed System Security Symposium(NDSS'04), San Diego, 2005.

[3] [APWG] *The Anti-phishing Working Group*, "Proposed Solutions to Address the Threat of E-mail Spoofing Scams," December 2003.

[4] [APJU] *The Anti-Phishing Working Group*, "Phishing Attacks Trend Report, June 2004".

[5] Microsoft Security Bulletin MS05-001, January 11, 2005.
http://www.microsoft.com/technet/security/bulletin/MS05-001.mspx

[6] Anti-Phishing Working Group, http://www.antiphishing.org.

[7] "Man sentenced for "botnet" attack on hospital," The Mercury News, August 25, 2006.
http://www.mercurynews.com/mld/mercurynews/news/local/states/california/northern_california/15364273.htm