DESIGN AND DEVELOPMENT OF A PLUG-IN FOR INTERNET BROWSERS TO PREVENT SPOOFING AND PHISHING ATTACKS

[1]NARAYAN YADAV, [2] VIKAS CHOUDHARY

[1]Research scholar, [2]HOD CSE
Institute of Engineering and Technology, Bhagwant University
Email: ydvnarayan@gmail.com

Abstract- Phishing is a way of attempting to personal information such as Username, Passwords (email), Credit card, and Bank etc. Faking an email or website is sometimes called a phishing attack. Phishing attacks involve setting up fake website or sending spam email in an attempt to lure potential victim to fake website. The "sender" field in an email can be changed easily and as long as the email message protocols are acceptable the message will be delivered. A phishing site can look just like the real one, with the same color schemes logout. A victim that attempts to use the site can unknown be submitting their personal data to phishes. Spoofing is the process of falsifying ones identity and masquerading as someone else. In this proposed work, induces analysis of different type of attacks, and try to find out res ion of usability, failures of browsers. Moreover it includes development and description of a security method by which a browser extension improves, using secure identification indicators. Users can assign a feed back to a secure site, presented by a tool bar when the browser presents that secure site; otherwise, the Tool bar presents the certified site's owner name, and the feedback of the Certificate Authority (CA) who identified the owner. Description and development of usability experiment, which measure, and prove the effectiveness, of security and identification indicators. Derive general secure-usability principles from my experiments and will investigate spoofing and phishing attacks and countermeasures, trying to protect naïve as well as expert users.

Keywords- Phishing, C4.5clustering, Phishing URL, Security, Internet Explore

I. INTRODUCTIONS

The web is the medium for an increasing amount of business and other sensitive transactions, for example for online banking and brokerage. Virtually all browsers and servers deploy the SSL/TLS (secure socket layer/transport layer security) protocols to address concerns about security. However, the current usage of SSL/TLS by browsers still allows web spoofing, (misleading users by impersonation or misrepresentation of identity or of credentials). Indeed, there is an alarming increase in the amount of real-life web-spoofing attacks, usually using simple techniques. Often, the swindlers (cheaters) lure the user to the spoofed web site, e.g. impersonating as financial institution, by sending her spoofed e-mail messages that link into the spoofed web-sites; this is often called a phishing attack. The goal of the attackers is often to obtain user-ID's, passwords/PINs and other personal and financial information, and abuse it e.g. for identity theft.

There are three main approaches to site identification indicators

1. Standard/classical indicators: the indicators available in typical current browsers, consisting mainly of the location (address/URL) bar, and of indicators of the activation of SSL/TLS (a padlock and the use of the protocol name httpsrather than http).

2.Certificate- derived identification indicator: presenting an identifier (feedback) for the site. If, as in current browsers, the identification is not always done by an entity trusted by the user (directly or by delegation), then we should also identify the entity responsible for the identification. Namely, in this case the identification indicator includes also a feedback for the Certificate Authority (CA), responsible for identifying the site.

3. User-customized identifiers:allowing users to choose a feedback for a securely identified site, and later presenting this feedback to identify this (SSL/TLS protected) site.

Mainly the proposed work concentrates on the attacks named as Phishing and Spoofing, which are described below:

1.1 Phishing:

Phishing is the process by which someone obtains private information through deceptive or illicit means in order to falsely assume another person's identity. The Phisher will use spoofed emails to lead the recipient to counterfeit websites. Once here, the victim is tricked into divulging credit card information, account usernames and passwords, social security numbers, etc.

1.1.1 Phishing techniques:

1. Spear Phishing

Phishing attempts directed at specific individuals or companies have been termed spear phishing. Attackers may gather personal information about their target to increase their probability of success.

2. Clone Phishing

A type of phishing attack whereby a legitimate, and previously delivered, email

containing an attachment or link has had its content and recipient address(es) taken and used to create an almost identical or cloned email. The attachment or Link within the email is replaced with a malicious version and then sent from an email address spoofed to appear to come from the original sender. It may claim to be a re-send of the original or an updated version to the original.

This technique could be used to pivot (indirectly) from a previously infected machine and gain a foothold on another machine, by exploiting the social trust associated with the inferred connection due to both parties receiving the original email.

3. Whaling

Several recent phishing attacks have been directed specifically at senior executives and other high profile targets within businesses, and the term whaling has been coined for these kinds of attacks.

1.1.2 Types of Phishing Attacks:

1. Phone Phishing:

Generally the hacker's a.k.a. Black Hats call some customers with automated calls wherein they mention that it is automated call from example bank, and it is mandatory verification,please enter your account number & pin etc.

2. Phishing by Website Forgery:

Web Forgery (also known as "Phishing") is a form of identity theft that occurs when a malicious Web site impersonates a legitimate one in order to trick you into giving up sensitive information such as passwords, account details, or credit card numbers. Phishing attacks usually come from email messages that attempt to lure the recipient into updating their personal information on fake, but very real looking, Web sites.

3. Tab-Nabbing

Tab-Nabbing is a popular phishing attack used by attackers to make users submit their login details like usernames and passwords by impersonating the most popular sites on web which no user doubts to be fake.

4. Phishing by Link Manipulation:

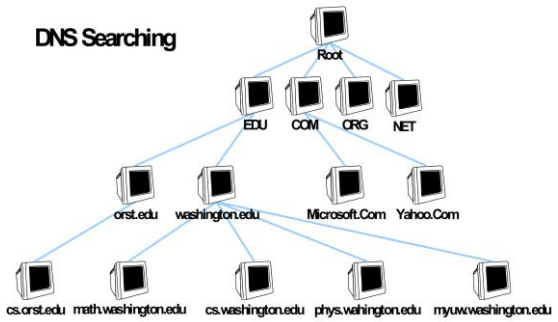
Link Manipulation is a phishing attack done mainly to miss-lead the user to a fake website or a "look-a-like" of some renowned site. The main trick used in this type of phishing is use of sub-domains. These are the technicalities which are not familiar to Non-I.T users and hence they are the primary targets of the black hats.

5. DNS-Based Phishing ("Pharming")

Pharming is the term given to hosts file modification or Domain Name System (DNS)-based phishing. With a pharming scheme, hackers tamper with a company's host files or domain name system so that requests for URLs or name service return a bogus address and subsequent communications are directed to a fake site. The result: users are unaware that the website where they are entering confidential information is controlled by hackers and is probably not even in the same country as the legitimate website.

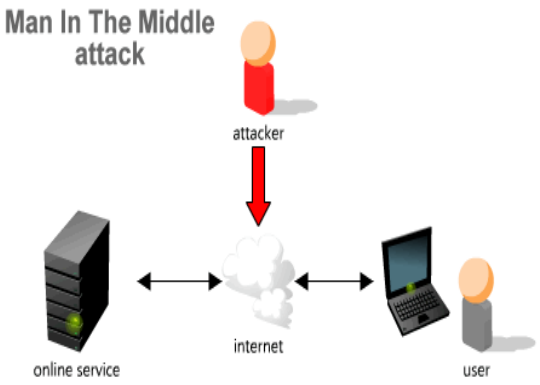
6. Phishing by Content-Injection:

It describes the situation where hackers replace part of the content of a legitimate site with false content designed to mislead or misdirect the user into giving up their confidential information to the hacker. For example, hackers may insert malicious code to log user's credentials or an overlay which can secretly collect information and deliver it to the hacker's phishing server.



7. Phishing by Man-in-the-Middle:

Man-in-the-Middle Phishing is harder to detect than many other forms of phishing. In these attacks hackers position themselves between the user and the legitimate website or system. They record the information being entered but continue to pass it on so that users' transactions are not affected. Later they can sell or use the information or credentials collected when the user is not active on the system.



8. Search Engine Phishing

Phishing search engine Occurs when phishers create websites with attractive (often too attractive)

sounding offers and have them indexed legitimately with search engines. Users find the sites in the normal course of searching for products or services and are fooled into giving up their information. For example, scammers have set up false banking sites offering lower credit costs or better interest rates than other banks. Victims who use these sites to save or make more from interest charges are encouraged to transfer existing accounts and deceived into giving up their details.

1.2Spoofing:

Spoofing is the creation of TCP/IP packets using somebody else's IP address. Routers use the "destination IP" address in order to forward packets through the Internet, but ignore the "source IP" address. That address is only used by the destination machine when it responds back to the source. A common misconception is that "IP spoofing" can be used to hide your IP address while surfing the Internet, chatting on-line, sending e-mail, and so forth. This is generally not true. Forging the source IP address causes the responses to be misdirected, meaning you cannot create a normal network connection. However, IP spoofing is an integral part of many network attacks that do not need to see responses (blind spoofing).

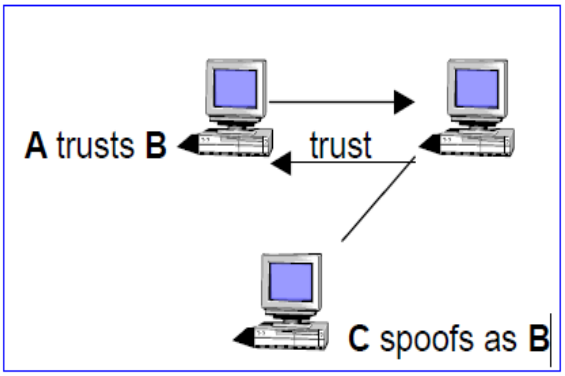
In this work as well as follow-up works done prior to the first publication of the current manuscript. The focus was on attacks and wary users. Specifically the users were accepted to correctly and carefully check indicators such as the URL (location) of the web site and the security lock (SSL/TLS) Indicator. These works showed clever Web Spoofing attacks using Scripts, Java applets or other features and bugs of common browsers to fool even naïve and expert users. The fixed Path phishing attack (method) process collected URLS, Classifies the fixed-path pattern of phishing URLS and in term of confidence level, prioritize the pattern sequence based on its length and volume.

Examples of spoofing:

Man-in-the-Middle
Packet sniffs on link between the two end points, and can therefore pretend to be one end of the connection
Routing Redirect
Redirects routing information from the original host to the hacker's host (this is another form of man-in-the-middle attack).
Source Routing
Redirects individual packets by hackers host
Blind Spoofing
Predicts responses from a host, allowing commands to be sent, but can't get immediate feedback.
Flooding
SYN flood fills up receive queue from random source addresses. Smurf/fraggle spoofs victims address, causing everyone respond to the victim.

1.2.1 Types of Spoofing

1. IP Spoofing:



IP Spoofing is a security exploit where an Intruder attempts to send packets to a system which appear to originate from a source other than the Intruder's own. If the target system already has an authenticated TCP session with another system on the same IP network, and it mistakenly accepts a spoofed IP packet, then it may be induced to execute commands in that packet, as though they came from the authenticated connection.

Improved reliability and routing filters in major Internet routers make this attack largely obsolete on the Internet in cases where theIntruder and target system are topologically dist

.

2. URL spoofing:

Another kind of spoofing is "webpage spoofing," also known as phishing. In this attack, a legitimate web page such as a bank's site is reproduced in "look and feel" on another server under control of the attacker. The main intent is to fool the users into thinking that they are connected to a trusted site, for instance to harvest user names and passwords. This attack is often performed with the aid of URL spoofing, which exploits web browser bugs in order to display incorrect URLs in the browsers location bar; in order

to direct the user away from the legitimate site and to the fake one. Once the user puts in their password, the attack-code reports a password error, then redirects the user back to the legitimate site.

3. Referrer spoofing:
Referrer spoofing or ref tar spoofing is the sending of incorrect referrer information in an HTTP request, sometimes with the aim of gaining unauthorized access to a web site. It is also used to improve the privacy of an individual using a web browser to view World Wide Web sites, by replacing valid referrer data with incorrect data, though most users simply suppress their web browser from sending referrer data, and may also modify other HTTP headers.

4. Caller ID spoofing:
In public telephone networks, it has for a long while been possible to find out who is calling you by looking at the Caller ID information that is transmitted with the call. There are technologies that transmit this information on landlines, on cell phones and also with VoIP. Unfortunately, there are now technologies (especially associated with VoIP) that allow callers to lie about their identity, and present false names and numbers, which could of course be used as a tool to defraud or harass. Because there are services and gateways that interconnect VoIP with other public phone networks, these false Caller IDs can be transmitted to any phone on the planet, which makes the whole Caller ID information now next to useless

5. E-mail Address Spoofing:
The sender information shown in e-mails (the "From" field) can be spoofed easily. This technique is commonly used by spammers to hide the origin of their e-mails and leads to problems such as misdirected bounces (i.e. e-mail spam backscatter).E-mail spoofing is a term used to describe (usually fraudulent) e-mail activity in which the sender address and other parts of the e-mail header are altered to appear as though the e-mail originated from a different source.By changing certain properties of the e-mail, such as the From, Return-Path and Reply-To fields (which can be found in the message header), ill-intentioned users can make the e-mail appear to be from someone other than the actual sender. The result is that, although the e-mail appears to come from the address indicated in from field (found in the e-mail headers), it actually comes from another source.

1.2.2. Types of Network spoofing:
1. Protocol Spoofing
In every network, there is a protocol group called the TCP. This protocol establishes, maintains and breaks down the connections. In the process of connecting the computer will send a check packet of data for verification. This adds up to the network traffic. This private network running over the public lines will incur extra charges as well. To avoid such a situation, the gateway can act as the remote computer and reply to the TCP messages. Here the network gateway is spoofing as the TCP connecting computer to reduce the traffic.

2. DNS spoofing
When a web page is requested through a web browser, it does not connect to the real web address. Before connecting with the web page, it will check with the Domain Name System to get the original IP address. But companies maintain their own DNS server to save response time. So when you click on the web page, you are connected to the in-house server and not to the public DNS server. Here DNS server is spoofing as the public DNS server of the company.

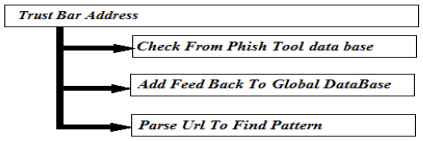
3. MAC spoofing
The entire device connected to a network will have a MAC address. When you register for internet connection, the internet service provider will register the MAC address for a more secured connection. Only the device with that MAC address can be connected to the network. If the user wants dual access points for internet, it will not be accepted. So the new device will send the information through the registered MAC address to gain access to the network by spoofing the registered device MAC address.

3.3 FACILITES REQUIRED FOR PROPOSED WORK:

1.NET Skills
C#, ASP.NET 2.0
2.Application & Web Servers
Internet Information Server, Base browser
3.RDBMS
Sql Server 2008
4.Operating Systems
MS Windows XP with service pack 3/MS Windows 7
5.Front End Tool
Visual Studio 2010
6.Other Skills:
Knowledge of previous security algorithms and threats
7.Other facilities:
Internet, books and study material on network security and .net framework.

3.4PROPOSED ALGORITHM:

The proposed architecture is shown below.



System Architecture

We will design a Trust Bar. This Trust bar address consists of three parts.

1.Check from phish tool data base.
2.Add feedback to global database.
3.Parse URL to find pattern.

To successfully implement the complete system required to involve the following steps:

1. Check from phish tank data base: We maintain a global database, this database keep phishing URL published by phish tank community. When a user enter a URL in internet explorer Trust bar fetch, this URL to match with the phishing database and find the answer URL is a phishing URL or not. When the URL found in phish tank database then show the message.
2. Add feedback to global database: We provide a feature in our Trust bar, "Add feedback". By using this user can also add web site feedback, sometimes a opened URL is not compatible for users and this URL harm the users' system then user add a feedback to stop again and again open this.
3. Parse URL to find pattern: In this phase we add a new concept to discover the phishing URL. By which we first find the pattern of phishing URLs' and then train an decision tree algorithm for that URLs' and after that supplied URL is classified using this trained algorithm.

Algorithms:

C4.5 builds decision tree from a set of training data in same way as ID3, using the concept of Information entropy.

The training data is a set P = P1, P2, P3, -----of already classified samples. Each sample Si=D1, D2, D3, ----- is a vector Where D1, D2, D3, ----- represent attribute or feature of the sample.

The training data is augmented with a vector P=P1, P2, P3, -----

Where P1, P2, P3----- represent the class to which each sample belongs.

The general algorithm for building decision trees is:-

1.Check for base cases
2.Let a_ best be the attribute with the highest normalized information gain
3.Create a decision node that splits on a best
4.Recourse on subsists obtained by splitting on a best, and add those nodes as children of node.

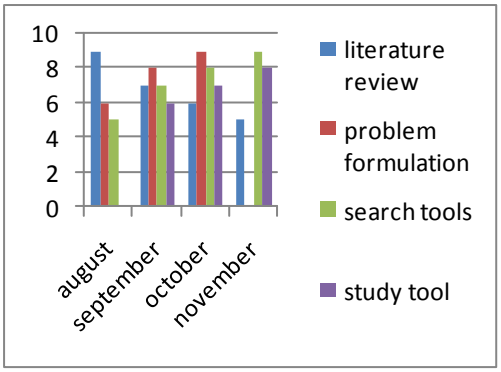
3.5 METHODOLOGY:

Research methodology is a way to systematically solve the research problem. It may be understood as a science of studying how research is done scientifically. In it we study the various steps that are generally adopted by a researcher in studying his research problem along with the logic behind them. Research methodology used in this research is experimental.

Research goes through following phases:-

1.the research problem.
2.Extensive literature survey.
3.Preparing the research design.
4.Determining sample design.
5.Collecting the data.
6.Formulating Selection of Tools applicable in the research work.
7.Learning the working of the tools and their applications in the research Analysis of data.
8.Creating a research proposal by the comparison of analysis of expected results and calculated results.
9.Execution of the proposed algorithm on the simulator.
10.Generalizations and interpretation of result.
11.Preparation of the report or presentation of the results, i.e., formal write-up of Conclusion reached

5.Work Plan With Timeline



REFERENCES

[1]. Amir Herzberg2 and Ahmad Jbara, (2006): Security and Identification Indicators Security and Identification Indicators for Browsers against Spoofing and Phishing Attacks"

[2]. ArunVishwanath,TejaswiniHerath b, Rui Chen c, Jingguo Wang d, H. RaghavRaoe,(2011):Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing modelsee front matter © Elsevier B.V. All rights reserved

[3]. CEAS 2009 -(July 16-17, 2009) Sixth Conference on Email and Anti-Spam :An Empirical Analysis of Phishing Blacklists. Mountain View, California USA

[4]. Cheng Hsin Hsu, Polo Wang, Samuel Pu(2007) :Identify Fixed-Path Phishing Attack by STC".

[5]. Colin Whittaker, Brian Ryner, MarriaNazif" (2007):Large-Scale Automatic Classification of Phishing Pages

[6]. Cormac Herley and DineiFlor˜encio Microsoft Research One Microsoft Way Redmond, "A Profitless Endeavour: Phishing as Tragedy of the Commons" WA,USAc.herley@ieee.org,dinei@microsoft.com

[7]. Divya James1 and Mintu Philip2,(2012):A NOVEL ANTI PHISHING FRAMEWORK BASED ON VISUAL CRYPTOGRAPHYInternational Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.1

[8]. Ian Fette, Norman Sadeh, Anthony Tomasic, (2007): Learning to Detect Phishing Emails, Banff,Alberta, Canada. ACM 978-1-59593-654-7/07/0005

a. Xun Dong, John, A.Clark, Jeremy L. Jacob (2010):Defending the weakest link: phishing websites detection by analyzing user behavior's, Published online: © Springer Science + Business Media, LLChttp;//www.quero.at

[9]. http://freefeast.info/general-it-articles/what-is-phishing-types-of-phishing-attacks explained-by-freefeast/

[10]. http://www.pcworld.com/article/135293/article.html

[11]. http://www.combofix.org/what-is-network-spoofing-and-what-are-the-types-of-it.php

[12]. http://www.sans.org/reading_room/whitepapers/threats/spoofing-overview- currentspoofing-threats_321

[13]. http://www.slideshare.net/Raza_Abidi/spoofing-techniques

[14]. http://en.wikipedia.org/wiki/Spoofing_attack