Volume 5, Issue 2, February 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

# Survey on Phishing Attacks

**S.S. Kulkarni, Mayank Tomar, Aastha Mittal, Sneha Arondekar, Aniket Nayakawadi**
I.T., SAE, Pune, Maharashtra,
India

***Abstract— Phishing attacks are one of the attacks which have become popular in recent times. A phishing attack attempts to acquire confidential and personal information of individuals or a firm. The most targeted domain of these phishing attacks to be known is financial domain, generally phishing attack attempts to obtain private information which can be exploited to gain access to bank account of the individual. Many solutions have been proposed ever since the phishing came into existence some validates the URL of the page to be visited to predict whether the page is legitimate or not while some recent practices in predicting the legitimacy of a pages different domains have been exploited such as visual cryptography. With increase in number of naive users of internet the chances of getting trapped in such attacks is quite a possible thing. So, it is necessary to provide some effective solution to this attack of phishing. This paper explains the details of various techniques used by phishing attacks to acquire sensitive information, it also provide advantages and disadvantages of the various anti-phishing technologies available at this moment to counter the phishing attacks.***

***Keywords— Phishing, Anti-phishing, Visual cryptography***

## I. INTRODUCTION

Phishing is an online identity theft which aims to acquire confidential information such as banking passwords and credit card details from users. Phishing attacks have in news in recent past because the volume of such attacks have increased drastically, according to a study 57 million US internet users have been identified as affected by phishing attacks and out of those 2 million were the actual victims of the attack and gave sensitive information to these attacks. Although these attacks have been in news for fairly long time but still the naive internet users became easy targets to such attacks just because of inexperience of the internet users. Attackers have been employing various technical spoofing tricks such as URL manipulation, hidden elements to look their site as similar as the target website. The most effective solution to phishing is educating users not to blindly follow links to web sites where they are to enter personal information such as passwords. However, expecting that all users will understand the phishing threat and think before clicking any link is unrealistic. There will always be users which will be tricked into visiting a phishing web site. Hence, it is essential for researchers and industry to provide solutions for the phishing threat. The fig 1 shows the domain wise target of the phishing attacks, it can be depicted from the image that the most popular domain among the phishing attacks in .com domain.

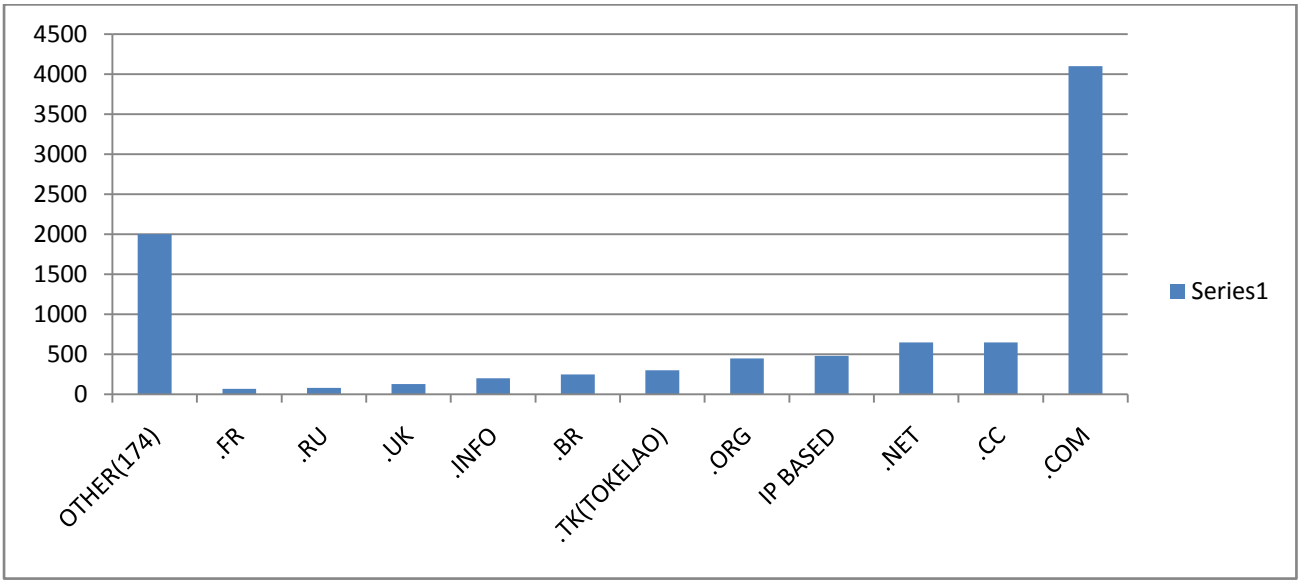


Fig. 1: Domain wise target of phishing attacks.

 

The fig2 shows which industry is most targeted among the targets of the phishing attacks. So this paper provides details of phishing and anti-phishing techniques implemented to counter the phishing attacks.

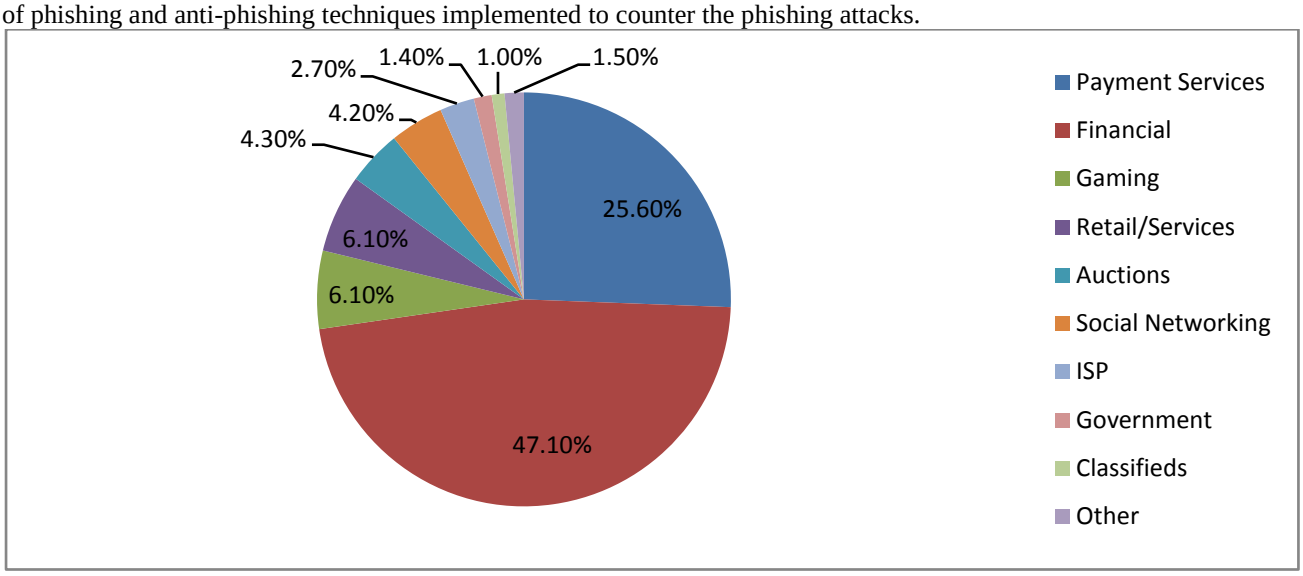


Fig.2. Most targeted Industry in 2011

## II. LITERATURE SURVEY

### *A. Phishing*

The act of sending an email to a user claiming to be a legitimate firm in an attempt to seek confidential data that can be exploited to perform identity theft is called as phishing.

Phishing email directs the user to visit a webpage which is identical to the phishing target webpage where they are asked to enter confidential personal information, such as a password, credit card, social security, or bank account numbers, that the legitimate organization already has. The webpage is however phished and set up only to access the information the user enters on the page.

### *B. Anti Phishing*

It is the act of detecting whether a visited webpage is legitimate or phished by extracting different properties of the webpage or by studying the visual contents of the webpage.

It is essential to provide an efficient technique to differentiate between a legitimate page and a phished page, to counter the ever increasing threat of phishing attacks.

### *C. Early Phishing on AOL*

Phishing on AOL(America online) has been the first ever appearance of phishing in the world of information technology , a software AOHell was released in early 1995, was a program designed to hack AOL users by allowing the attacker to gain access of various confidential details of the user. After AOL brought in measures in late 1995 to prevent using algorithmically generated PIN to open accounts, AOL crackers resorted to phishing for legitimate accounts and exploiting AOL.

In September 2003, the first known phishing attack against a retail bank was reported by The Banker in an article written by Kris Sangani titled Battle Against Identity Theft [1].By 2004, phishing was recognized as a fully industrialized part of the economy of crime: specializations emerged on a global scale that provided components for cash, which were assembled into finished attacks.

## III. CLASSIFICATION OF PHISHING ATTACKS

Phishing attacks can be classified into various types according to the way attack is done. According to many researchers the various types of phishing attacks has been described below.

### *A. Deceptive Phishing*

In this technique the phished webpage will ask the user to enter details to verify account information,, fictitious account charges, undesirable account changes, system failure requiring users to re-enter their information, new free services requiring quick action, and many other exciting offers so as to develop interest in users mind with the hope that the victim will click on the link as will provide the confidential personal information to the bogus webpage which can be further used to perform scams.

### *B. Malware Based Phishing*

This technique involves making run a malicious code on user's machine which is capable of performing tasks which will provide details of the confidential data entered by the user. Malware can be introduced in the user's machine as an attachment, by exploiting security vulnerabilities, as a downloadable file from a web site.

 

### *C. Web Trojans*

In this technique the pop-up invisibly runs when users are attempting to log in and they collect the personal information from the user's machine locally and transmits the information to the server the phisher is using to collect information of the victims.

### *D. System Reconfiguration Attacks:*

In this technique the phisher modify settings on a user's PC for performing various malicious operations without the knowledge of the user. For example: URLs in a favorites file can be altered to direct users to look a website which is visually identical to the target website. For example: a bank website URL may be changed from "www.gmail.com" to "www.gmaiL.com".

### *E. Pharming*

This technique modify the company's host file or DNS so that when the user want to log in or access that website , the changes made by the phisher will result in opening of phished website instead of the legitimate one. Hence used will submit the information to a phished page.

### *F. Content Injection Phishing*

In this technique the hacker replaces some part of code from the legitimate website which in turn results in submitting information to the server used by the phisher instead of submitting to the legitimate website.

### *G. Man-in-the-Middle Phishing*

In these attacks phisher positions 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.

## IV. CLASSIFICATION OF ANTI PHISHING TECHNIQUES

### *A. Content Filtering*

In this anti-phishing technique the emails before entering into the mailbox of the user is filtered using machine learning techniques such as Bayesian Additive Regression Trees (BART) or Support Vector Machines [2].

### *B. Blacklisting*

Blacklist is 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 [2].

### *C. Symptom Based Prevention*

Symptom-based prevention analyses the content of each Web page the user visits and generates phishing alerts according to the type and number of symptoms detected [2].

### *D. Domain Binding*

It is a client's browser based techniques where sensitive information (eg. name, password) is bind to particular domains. It warns the user when he visits a domain to which user credential is not bind.

## V. CONTENT BASED PHISHING

Gold Phish [3] tool implements this technique and uses Google as its search engine. This mechanism gives higher rank to well-established web sites. It has been observed that phishing web pages are active only for short period of time and therefore will acquire low rank during internet search and this becomes basis for content based anti-phishing approach [3]. The design approach can be broken down into three major steps. The first step is to capture an image of the current website in the user's web browser. The second step is to use optical character recognition techniques to convert the captured image into computer readable text. The third step is to input the converted text into a search engine to retrieve results and analyse the page rank.

**Advantages**: Generally Gold Phish does not result in false positive and provides zero day phishing [3].

**Disadvantages:** Gold Phish delays the rendering of a webpage. It is also vulnerable to attacks on Google's Page Rank algorithm and Google's search service [3].

## VI. CHARACTER BASED ANTI PHISHING TECHNIQUES

Many time phishers tries 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. <a href="URI"> Anchor text <\a> [4]

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.

Character based anti-phishing technique uses characteristics of hyperlink in order to detect phishing links. Link guard [6] is a tool that implements this technique. After analysing many phishing websites, the hyperlinks can be classified into

 

various categories as shown in fig 6. For detection of phishing sites Link Guard, first extracts the DNS names from the actual and the visual links and then compares the actual and visual DNS names, if these names are not the same, then it is phishing of category 1 If dotted decimal IP address is directly used in actual DNS, it is then a possible phishing attack of category 2[4].

If the actual link or the visual link is encoded (categories 3 and 4), then first the link is decoded and then analyzed. When there is no destination information (DNS name or dotted IP address) in the visual link then the hyperlink is analyzed. During analysis DNS name is searched in blacklist and white list .if it is present in white list then it is sure that the link is genuine and if link is present in blacklist then it is sure that link is phished one.

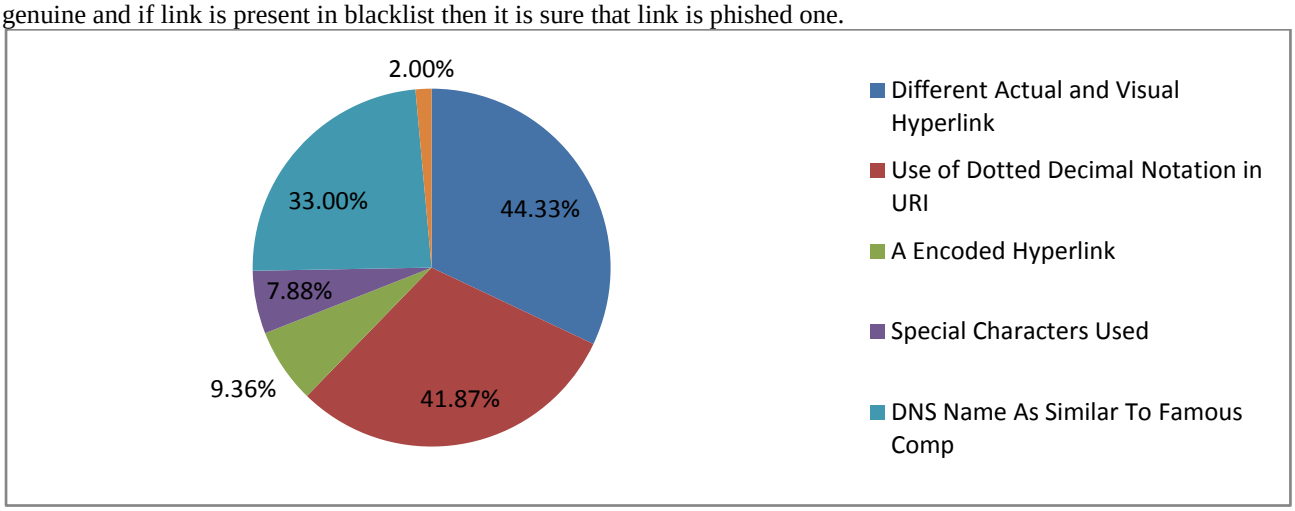


Fig.3 Phishing sites Link Guard

**Advantage**: it can not only detect known attacks, but also is effective to the unknown ones. Experiments showed that Link Guard, can detect up to 96% unknown phishing attacks in real-time [6]. For phishing attacks of category 1, it is sure that there are no false positives or false negatives .Link Guard handles categories 3 and 4 correctly since the encoded links are first decoded before further analysis [4].

**Disadvantage:** For category 2, Link Guard may result in false positives, since using dotted decimal IP addresses instead of domain names may be desirable in some special circumstances [4].

## VII. CONCLUSION AND FUTURE WORK

In the above study we can conclude that most of the anti-phishing techniques focus on contents of web age, URL and email. Character based anti-phishing techniques may result in false positive but content based approach never results in false positive. Attribute based approach consider almost all major areas vulnerable to phishing so it can be best anti-phishing approach that can detect known as well as unknown phishing attack. Identity based anti-phishing approach may fails if phisher gets physical access to client's computer.

As a future work on phishing we can do more work on server side security. In the server side security policy we use dual level of authentication for user by which only authentic user can get the access of his account, and to educate the user about this policy will results in avoiding user to give his sensitive information to phished web site.

## ACKNOWLEDGEMENT

We would like to thank our teacher Mr. S.S. Kulkarni for supporting us in making this paper and also by giving idea of how we can make the paper on literature survey.

## REFERENCES

[1] Sangani, Kris (September 2003). "The Battle Against Identity Theft". *The Banker* **70** (9): 53–54.

[2] Hicham Tout, William Hafner "Phishpin: An identity-based anti-phishing approach" in proceedings of international conference on computational science and engineering, Vancouver, BC, pages 347-352, 2009.

[3] Michael Atighetchi, Partha Pal "Attribute-based prevention of phishing attacks" Eighth IEEE international symposium on network computing and application, 2009.

[4] Juan Chen, Chuanxiong Guo-"Online Detection and Prevention of Phishing Attacks (Invited Paper)"in proceedings of Communicational and networking in china, first international conference, Beizing, pages 1-7, 2007.

[5] Matthew Dunlop, Stephen Groat, and David Shelly" GoldPhish: Using Images for Content-Based Phishing Analysis", in proceedings of internet monitoring and protection(ICIMP),fifth international conference, Barcelona, Pages 123-128, 2010.

[6] Archit Shukla, Lalit Gehlod, " A survey on phishing detection and prevention technique,"International Journal Of Engineering And Computer Science ISSN: 2319-7242 Volume 3 Issue 5 may, 2014 Page No.6255-6259.

 