ISSN: 2321-7782 (Online)

Volume 3, Issue 7, July 2015

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Survey Paper / Case Study

Available online at: www.ijarcsms.com

An Overview of Methods to Detect Phishing Web Pages

K. Selvan[1]
Research Scholar and Assistant Professor
Department of Computer Science
JJ College of Arts and Science, Pudukkottai.TN

Dr. M. Vanitha[2]
Assistant Professor
Department of Computer Science
JJ College of Arts and Science, Pudukkottai.TN

Abstract: Phishing is an electronic online theft that deals with social engineering methodology to illegally acquire and use someone else's data on behalf of legitimate website for own easy profit (e.g. Steal of user's password and credit card details). It is affecting all the major sectors of industry day by day with a lot of misuse of user credentials. To protect users against phishing, various anti-phishing techniques have been proposed that follows different strategies like client side and server side protection. In this paper we 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. This paper deals various anti phishing techniques and also deals analyzing and categorizing these methods.

Keywords- Anti-phishing, Identity theft, Phishing, Mutual authentication.

I. INTRODUCTION

WWW is the repository of electrically stored information and online services simplify our lives. Web browsers and servers take almost enough care to make guarantee the safe business transactions. Still they are vulnerable to attacks such as phishing. Interacting with an online service requires a certain degree of technical sophistication that not all internet users possess. Phishing is a wider spread pernicious practice and it is a form of online fraudulent activity in which an attacker aims to steal a victims sensitive information, such as an online banking password or a credit card number aiming to make an easy profit by means of financial transaction. Victims are tricked into proving such information by a combination of spoofing techniques and social engineering. In practice the victims often receive an email/SMS/phone call that tries to convince them to visit a web page that has been prepared by attacker. Phishing pages have to quite similar to the legitimate web pages in order to deceive users. For this reason phishes normally use a technique called visual deception. Phishing attackers use various tactics and convince naïve users to visit bogus sites. Ordinary internet users cannot become familiar with all these phishing techniques easily. The phishing web site usually looks exactly like a known legitimate website, mimics logos, images and text information. Phishing is a model problem for illustrating usability concerns of privacy and security. Phishing attacks are growing at a torrid pace, has a huge negative impact on service providers and consumers revenues, business relationships, marketing efforts and confidence on internet. In this paper we are going to address the existing anti-phishing mechanism and provide a comprehensive up to date survey of these methods.

II. ANALYSSIS OF PHISHING APPROACHES AND METHODS

The goal of this paper is to present and to understand the existing approaches, techniques to spoof naïve internet users and also the same to detect phishing sites. Currently, there are two main approaches used to detect phishing sites, URL black lists and page analysis.

URL black lists:

Black list holds URLs that refer to sites that are considered malicious. Whenever a browser loads a page, it queries the black lists to determine, whether the currently visited URL is considered legitimate. The Black can be stored locally at the client or hosted at a central server. `an important factor for the effectiveness of a blacklist is its coverage. The coverage indicates how many phishing pages on the www are included in the list. `The quality indicates how many non-phishing sites are incorrectly included in the list and how many escaped from the list.

Page Analysis

Page analysis techniques examine properties of the web page and the URL to distinguish phishing and legitimate sites. Page properties are typically derived from the pages HTML source code. The main properties are the number of password fields, the no of links, browser indicators tool bars, windows, address bar, and status bar and domain name.

III. List The Strategies Organized Along Three Dimensions

Lack of knowledge on computers, web visual deception, and lack of attention regarding security.

1. **Lack of knowledge**: Many users lack the underlying knowledge of how operating systems, application software's, email and the web works together to offer secure data exchange between client and server. Phishing attackers effectively use this in several ways. Domain name, email header, a closed padlock icon in the browser, SSL certificates are the leading concepts under this.

2. **Visual deception**: Phishes use visual deception tricks to mimic legitimate text, images, and windows. Even users having enough knowledge on computer and internet may be deceived by visually deceptive text, windows masking underlying windows and deceptive look and feel.

3. **Bounded attention**: Due to the lack of knowledge on web servers and its functioning even an experienced users can be spoofed by lack of attention to security indicators and its absence in the web page, we are visiting. Bounded attention also mean , unaware of correct business transactions related guidelines given by service providers, government , social welfare and research institutes.

In this paper we present different existing strategies for predicting and confirming phishing pages. *An anti phishing strategy based on visual similarity assessment [1]* : Proposed an anti phishing strategy uses visual characteristics to identify potential phishing sites measures suspicious pages similarity to actual legitimate sites. The first of two sequential processes in the site watcher system runs on local servers and monitors emails for keywords and suspicious URLs. The second process then compares the potential phishing pages against actual pages and assesses visual similarities between them in terms of key regions, page layouts, and overall styles .phishing page detection should focus more on visual similarities not only text features, full page similarity is necessary is the negative comment. *Counteracting phishing page polymorphism [2]: an image layout analysis approach.* In this approach during the detection process , it take the whole page as an image instead of analyzing the HTML code. The layout similarity between an authentic page and a suspect page is analyzed. Based on the similarity score, then determine whether a page is phishing or not. Both the authentic page and the suspect page as images and apply Otsu's thresholding method to transform them into black and white images and examine the pixels of the images, take all adjacent pixels with identical colors as a blob, and record their size and location information. Compare the sizes and locations of all block pairs, differences in the location, width, and height of a pair are smaller than certain thresholds, and degree of similarity is calculated to confirm the suspect page is phishing. *A Hybrid Authentication Mechanism for Preventing Phishing Attacks On E-banking Systems [3]:* Here a technique combination fingerprint recognition and sitekey authentication for securing e-banking system is introduced to tackle phishing attacks by establishing the authenticity of protected web sites and the users. It was designed to prevent unauthorized access to a web site even the case of identity theft by asking the web site user questions with difficult-to-guess answers. *EyeBit: Eye Tracking approach for Enforcing Phishing Prevention Habits [4].* Many participants based studies on the habit of secure web browsing and to understand decision patterns of end users while the fundamental problem in phishing is the fact that they are deceived. Checking the browsers address bar is beneficial for the end users to be aware of phishing. The reader should note that modern web browsers do show the website's URL and security information in the address bar. Even the knowledge of URL and security are also important for phishing prevention and are the strong motivation for seeing there, the both of them could not work before the users did not see the bar. *The Battle against Phishing: Dynamic Security Skins[5].* The author presents a technique for a user to distinguish authenticated web pages from insecure web pages. Due to the web is a client server based application the remote server generates an abstract image that is unique for each user and each transaction. This image is used create a skin, which customizes the appearance of the server's webpage. The browser computes the image that it expects to receive from the server and displays it in the users trusted widow. To authenticate content from the server, the user can visually verify the images match. Secure Remote Password Protocol with SSL is implemented. *Prevention of Phishing Attacks Based On Discriminative key point Features of WebPages.[6]* This approach states that, the page is stored as an image then the next step is to get the original page snapshot which is also saved as an image. Then select the source image as well as the target image compared. The images are compared by using color ratio. The difference is noted and reported to the user. When the difference is zero then the page is not a phishing page. This anti phishing tool is very efficient because it compares the phishing and authentic pages based on the visual appearance level, instead of rather than using text based analysis. *A Framework for Predicting Phishing Websites Using Neural Networks.* There is a big list of phishing indicators. The developed neural network for predicting website estimates ranges of phishing indicators genuine, doubtful, and legitimate and using these values rules will be formed and the network will be trained to give output that ranges between very legitimate, legitimate, suspicious, phish and very phi shy.

IV. Conclusion

In the above study we can conclude that most of the anti phishing techniques focus on contents of web page, URL and email. Character based anti-phishing approach may result in false positive but content based approach never results in false positive. Page analysis 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. Black list 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 their account, and to educate the user about this policy will results in avoiding user to give his sensitive information to phished web site.

References

1. An antiphishing strategy based on visual similarity assessmentInternet Computing, IEEE ...> Volume:10 Issue:2 Wenyin Liu ; Dept. of Comput. Sci., City Univ. of Hong Kong, Hong Kong ; Xiaotie Deng ; Guanglin Huang ; Fu, A.Y.
2. Counteracting Phishing Page Polymorphism:An Image Layout Analysis Approach Ieng-Fat Lam1, Wei-Cheng Xiao1, Szu-Chi Wang2, and Kuan-Ta Chen1 1 Institute of Information Science, Academia Sinica 2 Institute of Computer Science and Information Engineering, www.ijetae.com/files/Volume4Issue12/IJETAE_1214_101.pdf
3. A Hybrid Authentication Mechanism for Preventing Phishing. Attacks on E-banking Systems: The Nigeria Case Study. G. B Ogunniye1, O. M Afolabi2. www.necoma-project.eu/m/filer.../d2/.../miyamoto-badgers2014.pdf
4. EyeBit: Eye-Tracking Approach for Enforcing. Phishing Prevention Habits. Daisuke Miyamoto. ∗. , Takuji Iimura. ∗. , Gregory Blanc. †. , Hajime Tazak
5. www.eecs.berkeley.edu/~tygar/.../Phishing/Battle_against_phishing.pdfby R Dhamija - Cited by 482 - Related articles The Battle Against Phishing: Dynamic Security Skins. Rachna Dhamija. University of California, Berkeley rachna@sims.berkeley.edu. J.D. Tygar. University
6. Predicting Phishing Websites using Neural Network trained with Back- ... ArtificialNeural Network (ANN)
7. citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.227.9309...by M Rajalingam - 2012 - Cited by 8 - Related articlesAttempts to stop phishing by preventing a user from interacting with a malicious web site have ... based on discriminative key point features in WebPages. weblidi.info.unlp.edu.ar/WorldComp2013-Mirror/p2013/ICA2183.p April through June 2014 saw the second-highest number of phishing sites

AUTHOR(S) PROFILE

K.Selvan completed his under graduation in Arul Anandhar College, Karumathur, Madurai Kamaraj University during 1992- 1995 and his post graduation in St.Joseph's College ,Trichy ,Bharathidhasan University during 1995-1998. Now he is doing his research work in JJ College of Arts and Science(Autonomous).Pudukkottai, Bharathidhasan University in Network Security area. He has 17 years of teaching experience in Arts and Science colleges. He guides many M.Phil., research scholar and delivered special lectures in various topics related to computer science.