Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 2, February 2014, pg.739 – 744

RESEARCH ARTICLE

AN ANTI-PHISHING FRAMEWORK WITH NEW VALIDATION SCHEME USING VISUAL CRYPTOGRAPHY

Mr. Bhushan Yenurkar[1], Mr. Shrikant Zade[2]

[1]Student, M.Tech III semester, CSE Department, Priyadarshini Institute of Engineering and Technology, Nagpur (M.S)

bhushanyenurkar@yahoo.in

[2]Assistant Professor, CSE Department, Priyadarshini Institute of Engineering and Technology, Nagpur (M.S)

ABSTRACT

Phishing is a kind of online security attack where the attacker creates a replica of an existing web page to fool users in order to hack their personal, financial, or password data. Phishing is a form of online fraudulent activity in which an attacker aims to steal a victim's sensitive information, such as an online banking password or a credit card number. Victims are tricked into providing such information by a combination of spoofing techniques and social engineering. In this paper we have proposed a new approach named as "Anti phishing framework with interactive captcha validation scheme using visual cryptography" to solve the problem of phishing. It uses visual cryptographic schemes to counter phishing pages where one secret captcha image share resides with user and the other secret shares reside in server. During authentication a genuine server forwards its share and the user forwards his share resulting in a secured access to the system via a reconstructed captcha. But the traditional captcha is prone to character recognization attacks and third-party human attacks. To overcome this problem we used here new generation of captcha which is known as interactive captcha to counter the both attacks. By recording CAPTCHA solving time on a per-character basis, we propose to use Detection Threshold Algorithms for CAPTCHA that enables a server to detect and reject third-party human attacks in ways not possible with existing CAPTCHAs. Combined with visual cryptographic schemes, we offer a dynamic or interactive captcha that can thwart all possible authentication threats.

Keywords: Phishing, Visual Cryptography, Image Captcha, Security

1. INTRODUCTION

Phishing is an attempt by an individual or a group to get confidential information such as passwords and credit card information from unsuspecting victims for identity theft, financial gain and other fraudulent activities. Phishing is an act of attempting to acquire sensitive information of a person by masquerading as a trust worthy entity in electronic transaction. Phishing is typically carried out by e-mail spoofing or instant messaging. Phishing e-mails contain links to websites infected with malware. Phishing is generally carried out through e-mail spoofing and by mimicking the web ages of original websites which look exactly the original ones. Here, attacker sends a mail to the person whose details he wants to track. In the mail attacker hides his true identity and generally he sends a link which appears similar to the genuine website like bank website etc.., Here, attacker adds some message to mislead the user. Innocent users think it is true and they login to the site providing their credentials and thus falling prey for Phishing attack. So here introduces a new and secure method which can be used to prevent phishing attacks which is named as "An Anti-phishing framework with interactive captcha validation scheme using visual cryptography". In this method, we provide a provision to the user to check whether the website he is willing to visit is a genuine website or a phishing website. So, by knowing these he can securely perform his further proceedings or transactions.

Here, we used the concept of an improved visual cryptography. Visual Cryptography (VC) is used here to divide the image captcha into shares and in order to reveal the original image captcha appropriate number of shares should be combined.

1.1 Visual Cryptography:-

Visual cryptography schemes were independently introduced by Shamir and Blakley, and their original motivation was to safeguard cryptographic keys from loss. These schemes also have been widely employed in the construction of several types of cryptographic protocols and consequently, they have many applications in different areas such as access control, opening a bank vault, opening a safety deposit box, or even launching of missiles. A segment-based visual cryptography suggested by Borchert can be used only to encrypt the messages containing symbols, especially numbers like bank account number, amount etc. The VCS proposed by Wei-Qi Yan et al., can be applied only for printed text or image. Naor and Shamir introduced the visual cryptography scheme (VCS) as a simple and secure way to allow the secret sharing of images without any cryptographic computations.

Most of the previous research work on VC focused on improving two parameters: pixel expansion and contrast. In these cases all participants who hold shares are assumed to be honest, that is, they will not present false or fake shares during the phase of recovering the secret image. Thus, the image shown on the stacking of shares is considered as the real secrete image. But, this may not be true always. So cheating prevention methodologies are introduced by Horng et al., and Hu et al.,. But, it is observed in all these methodologies, there is no facility of authentication testing.

VCS is a cryptographic technique that allows for the encryption of visual information such that decryption can be performed using the human visual system. We can achieve this by one of the following access structure schemes.

1.(2, 2)- Threshold VCS scheme- This is a simplest threshold scheme that takes a secret message and encrypts it in two different shares that reveal the secret image when they are overlaid.

2. (n, n) -Threshold VCS scheme-This scheme encrypts the secret image to n shares such that when all n of the shares are combined will the secret image be revealed.

3.(k, n) Threshold VCS scheme- This scheme encrypts the secret image to n shares such that when any group of at least k shares are overlaid the secret image will be revealed.

In the case of (2, 2) VCS, each pixel P in the original image is encrypted into two sub pixels called shares. Note that the choice of shares for a white and black pixel is randomly determined (there are two choices available for each pixel). Neither share provides any clue about the original pixel since different pixels in the secret image will be encrypted using independent random choices. When the two shares are superimposed, the value of the original pixel P can be determined. If P is a black pixel, we get two black sub pixels; if it is a white pixel, we get one black sub pixel and one white sub pixel.

2. LITERATURE SURVEY

There have been many authentication methods that have been proposed by researchers for authentication. Some of the most prominent of them have been discussed here.

Initially there have been some techniques where user based mechanisms are used to authenticate server. Automated Challenge Response Method (ARM) [1] is one such authentication mechanisms where challenge generation module in server requests for response from Challenge-Response interface in client. Then Challenge-Response module calls get response application installed in client machine. Once this is done, user credentials are demanded from client and it is

validated by server and thus transaction is made secure. This ensures two way authentication and also prevents man-in-middle attacks as response is obtained from executable which is called by browser and third man cannot interrupt at any cost.
There are also some Domain Name Service (DNS) based anti-phishing approaches[2] techniques which mainly include blacklists, heuristic detection, the page similarity assessment etc.., But, there are many disadvantages with these approaches.

Blacklist based technique is a DNS based anti-phishing approach commonly used by browsers. Some Work Groups provide an open blacklist query interface. Some of the most used browsers like Netscape Browser8.1, Google Safe Browsing (a feature in Google Toolbar for Firefox), Internet Explorer7 use blacklists to protect users when they are browsing through Internet. Blacklists are lists of URLs of some of the phishing sites.
There are many shortcomings in this approach. This technique has low false alarm probability, but it cannot detect the websites that are not in blacklists. Life cycle of phishing websites is too short for establishment of blacklists which makes this technique inaccurate.

Heuristic-based anti-phishing technique is a technique where a webpage is checked to find out whether the page has any of the phishing heuristics characters like host name, checking URL for common spoofing techniques and checking against previously seen images.
This method does not yield accurate results as even the attackers are aware of such techniques and they use some strategies so that they are not detected. So some *similarity assessment methods* have been proposed to detect phishing websites. For example, CATINA[4] is a content similarity based approach to detect phishing websites. Here, initially calculates the suspicious page's lexical signature using TF-IDF and then feeds this to search engine. Basing on the suspicious page's sort order in the search results the site is checked for its legitimacy.

There are many other similarity based assessment methods. Some of them are mentioned here.
Liu Wenyin and Anthony Y. Fu etc. [5] [6] proposed a page visual similarity assessment method to detect phishing websites, if a web page is similar to a financial organization's page, but it is not the organization's web page itself, it is considered a phishing site's page. JungMin Kang and DoHoon Lee [7] proposed the URL similarity assessment method, if an URL is similar to a bank's URL, but it is not the bank's URL, it is considered a phishing website's URL. There is low assess accuracy rate for the URL and content similarity assessment techniques. The speed of calculating the visual similarity between pages is too slow, so it is only used for phishing-spam detection generally.

Another scheme named A Three-Factor Authentication Scheme named *Phish-Secure* has been proposed to counter phishing[8].

As a first factor of authentication, an image similarity detection is done which helps in finding out which page the user tends to visit, then it is checked for Phishing. For this purpose a system captures the image of a webpage in a particular resolution in the required format. This image is termed as Visual image. If the attacker is going to create a Phishing site he is going to use the replica of the original webpage in order to fool the users. Now Phish-Secure gets the Visual image of the visited page and collects the mean RGB value of the image. This is termed as M_RGB. The database with Phish-Secure uses consists of details about the page which has to be authenticated. The actual mean RGB of various web pages is stored in the database which is denoted as AM_RGB. Phish-Secure will utilize this information and make a comparison to find out the similarity between the visited page and the page in the database. The similarity is obtained in means of percentage, if the percentage of similarity (PS) is greater than 99 % then Phish-Secure concludes which website the user is tending to visit. This is carried out by taking the corresponding URL in the database and checking is done in order to find whether the site is Phishing or not.

As a second factor of authentication Phish-Secure grabs the destination IP in Layer 3 which gives information about to which IP address the user is getting connected, this is referred as C_IP. If an attacker's web server IP address has already been found guilty the particular IP is blacklisted. Phish-Secure check this Blacklist with the C_IP and will warn the user. On the other hand if the C_IP is not found in Blacklist, further verification is done in the following step.
Here in this step Phish-Secure grabs the actual list of IP address of the provider which he tends to connect. This is because any provider may have multiple servers for the purpose of load balancing and the user may be connected to his location accordingly.

In order to avoid any confusion Phish-Secure gets the list of IP address which is referred to as actual IP and is checked with the C_IP (i.e.) the IP address to which the user is getting connected. If these two IP address are same Phish-Secure identifies the particular site as genuine and returns a message as authenticated. On the other hand if there is a mismatch in the above verification Phish-Secure identifies the site as Phishing and warns the user. In addition to this the C_IP is added to the black list so that in future if the attacker uses the same web server and tries to attack, Phish-Secure detects the site as Phishing in the second step.

To provide the user traffic as users manage more accounts, OpenID was proposed. OpenID provides single sign-on (SSO) service, that is, we can enjoy service of multiple sites by signing in only once. But this is vulnerable to phishing attack, So many methods have been proposed to overcome this drawback. Some of them are mentioned here.
"New Anti-Phishing Method with Two Types of Passwords in OpenID System"[13], is one such method. In this method, two types of passwords have been put forward for anti-phishing for OpenID users. In this method only one fixed passwords and many temporary (session) passwords are used. Fixed passwords are bound to a PC or any electronic device which user owns or which he frequently uses. Temporary passwords are used when user logs in different systems, for this user is sent temporary passwords to his mobile or mailbox. This method effectively avoids phishing.

Haijun Zhang, Gang Liu, Tommy W. S. Chow [10] proposed a textual and visual content based antiphishing mechanism using Bayesian approach. This framework synthesizes multiple cues, i.e., textual content and visual content, from the given web page and automatically reports a phishing web page by using a text classifier, an image classifier, and a data fusion process of the classifiers. A Bayesian model is proposed to estimate the threshold, which is required in classifiers to determine the class of web page. It also develop a Bayesian approach to integrate the classification results from the textual and visual contents. The main contributions of this paper are threefold. First, it proposes a text classifier using the naive Bayes rule for phishing detection. Second, it propose a Bayesian approach to estimate the threshold for either the text classifier or the image classifier such that classifiers enable to label a given web page as "phishing" or "normal." Third, a novel Bayesian approach to fuse the classification results from the text classifier and the image classifier is proposed.

There are various mutual authentication methods using cell phones such as browsing using phones, password generation etc. But, there are various problems of these methods which are discussed and some of them having their own advantages and disadvantages.

3. PROPOSED RESEARCH METHODOLOGY

In order to prevent phishing attacks, we are proposing a new methodology which is helpful to detect the phishing website. Our methodology helps the user to protect their password and other confidential and sensitive information from the phishing websites.
In our proposed methodology, we are using here multiple phases for authentication of a genuine user and also prevent user from phishing attack.

In the registration phase, we just simply asked the user to enter their username and password for the secure website. The password can be alphanumerical to provide more secure environment. After entering the username and password then the user must be given an image captcha. The user is asked to enter the text shown in image captcha. Then, after submitting all of the above details the image captcha is divided into two shares using visual cryptography i.e user's share and server share. The user can download it's share and kept with them. As soon as at a time of downloading user's share the server share is automatically sent to the any confidential or secure server. The original image captcha is also sent to the user along with their share which can be used further for verification during login phase. The original image captcha is also sent to the server as confidential data. After the completion of registration, the user can change the password when it is needed.

In the Login phase, the user is first prompted for the username (user id).Then the user is asked to upload his share which is kept with him. After uploading his share, this share is sent to the server where the user's share and share which is stored in a confidential server, for each user, are concatenated together to produce the image captcha. The image captcha which is generated is shown to the user. Here the end user can check whether the displayed image captcha matches with the captcha created at the time of registration. After that here we are proposing a new method which protects the user and

captcha from image based attacks and third party human based attacks. This method is generally a captcha solving test. The test starts with by clicking on the captcha image which is shown to the user. Then, the user is asked to follow certain sequences in order to successfully solve this captcha test. After successfully completion of this sequences, then finally the user is requested to enter their password. Using this, the user can log in into the website and can securely perform further proceedings. Using the username and image captcha generated by stacking two shares one can verify whether the website is genuine/secure website or a phishing website and can also verify whether the user is a human user or not.

In our proposed system, our system successfully detects all type of phishing attacks and CAPTCHA plays an important role in protecting Internet resources from attacks by automated scripts. Because, many researchers proposed that the current existing captcha can be attacked by using various well-known techniques and within submission time it can be known to the attackers. So, our proposed systems used visual cryptographic schemes and newly designed captcha to counter phishing pages where one secret captcha image share resides with user and the other secret share resides in server. During authentication a genuine server forwards its share and the user forwards his share resulting in a secured access to the system via our newly designed image captcha.

4. CONCLUSION

In this paper, we have proposed a system in which CAPTCHA plays an important role in protecting Internet resources from attacks by automated scripts. Prior systems used visual cryptographic schemes to counter phishing pages where one secret captcha image share resides with user and the other secret share resides in server. During authentication a genuine server forwards its share and the user forwards his share resulting in a secured access to the system via an image captcha. The image captcha always happens to be same and is prone to character recognition based attacks. To solve this problem we propose to use visual cryptographic schemes to counter phishing and an interactive captcha to counter character recognition based attacks. By recording CAPTCHA solving time on a per-character basis, we use detection threshold algorithms for CAPTCHA that enable it to detect and reject 3rd party human attacks in ways not possible with existing CAPTCHAs. In our system, we offer dynamic captcha that can thwart all possible authentication threats.

References

[1] Thiyagarajan, P.; Venkatesan, V.P.; Aghila, G.; "Anti-Phishing Technique using Automated Challenge Response Method'", in Proceedings of IEEE- International Conference on Communications and Computational Intelligience, 2010.

[2] Sun Bin.; Wen Qiaoyan.; Liang Xiaoying.; "A DNS based Anti-Phishing Approach," in Proceedings of IEEE- Second International Conference on Networks Security, Wireless Communications and Trusted Computing, 2010

[3] Nourian, A.; Ishtiaq, S.; Maheswaran, M.;" CASTLE: A social framework for collaborative antiphishing databases", in Proceedings of IEEE- 5th International Conference on Collaborative Computing:Networking, Applications and Worksharing, 2009.

[4] Sid Stamm, Zulfikar Ramzan, "Drive-By Pharming", v4861 LNCS,p495-506, 2007, Information and Communications Security - 9th International Conference, ICICS 2007, Proceedings.

[5] Anthony Y. Fu, Liu Wenyin, "Detecting Phishing Web Pages with Visual Similarity Assessment Based on Earth Mover's Distance (EMD)",IEEE Transactions on Dependable and Secure Computing, v 3, n 4, p301-311, October/December 2006.

[6] Wenyin Liu, Xiaotie Deng, Guanglin Huang, and Anthony Y. Fu, "An Antiphishing Strategy Based on Visual Similarity Assessment", IEEE Internet Computing, v 10, n 2, p 58-65, March/April 2006.

[7] JungMin Kang, DoHoon Lee, "Advanced White List Approach for Preventing Access to Phishing Sites", 2007 International Conference onConvergence Information Technology, ICCIT 2007, p 491-496, 2007

[8] Nirmal, K.; Ewards, S.E.V.; Geetha, K.; "Maximizing online security by providing a 3 factor authentication system to counter-attack 'Phishing'", in Proceedings of IEEE- International Conference on Emerging Trends in Robotics and Communication Technologies, 2010.

[9] Qingxiang Feng.; Kuo-Kun Tseng.; Jeng-Shyang Pan.; Peng Cheng and Charles Chen.; "New Antiphishing Method with Two Types of Passwords in OpenID System", in Proceedings of IEEE Fifth International Conference on Genetic and Evolutionary Computing,2011.

[10] Haijun Zhang , Gang Liu, and Tommy W. S. Chow, "Textual and Visual Content-Based Anti-Phishing:A Bayesian Approach," *IEEE Trans. Neural Netw.*, vol. 22, no. 10, pp. 1532–1546, Oct. 2011.

[11] M. Naor and A. Shamir, "Visual cryptography," in Proc. EUROCRYPT, 1994, pp. 1–12.

[12] Divya James and Mintu Philip, "A Novel Anti Phishing Framework Based on Visual Cryptography", International Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.1, January 2012.

[13] R. Dhamija and A. Perrig, "Deja Vu: A User Study Using Images for Authentication," in Proceedings of 9th USENIX Security Symposium, 2000.

[14] Wen-Pinn Fang, "Visual Cryptography in reversible style,"IEEE Proceeding on the Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing(IIHMSP2007), Kaohsiung, Taiwan, R.O.C, 2007, 11, 26～2007, 11, 28.

[15] JungMin Kang, DoHoon Lee, "Advanced White List Approach for Preventing Access to Phishing Sites", 2007 International Conference on Convergence Information Technology, ICCIT 2007, p 491-496, 2007

