IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 2, Issue 3, June-July, 2014
ISSN: 2320 – 8791 (Impact Factor: 1.479)
www.ijreat.org

An Anti-Phishing Framework With Interactive Captcha Validation Scheme Using Visual Cryptography

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

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

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

ABSTRACT

Phishing is identified as a major security threat on internet today. Phishing is a kind of online fraudulent activity in which an attacker aims to steal a user's personal or sensitive information such as an online banking password or a credit card information. The attackers tricked the user's into providing their personal information by using some spoofing techniques and social engineering techniques. 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. Our proposed methodology uses visual cryptography scheme and new captcha scheme to counter attack phishing. The visual cryptography scheme is used here to divide the captcha image into two different shares such that one of the share kept by the user and another share is forwarded to the 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 some researchers proposed that the traditional captcha is prone to character recognization attacks and third-party human attacks. In order to overcome this problem we used here new generation of captcha which is known as interactive captcha to counter the both attacks. In this new captcha scheme, we record the CAPTCHA solving time on a per-character basis, which enables a server to detect and reject third-party human attacks in ways not possible with current CAPTCHAs. By combining this new captcha scheme with visual cryptography scheme, we provide 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 a group or an individual to get user's confidential information such as passwords and credit card information from unsuspecting victims for their financial gain, identity theft and other fraudulent activities. Phishing is attempt of acquiring user's sensitive information by masquerading as a trust worthy entity in electronic transaction. Phishing is generally carried out by e-mail spoofing or instant messaging. In e-mail spoofing, phishing e-mails contain the links to the websites infected with malware. In another way, phishing is generally carried out by mimicking the web pages of original websites which look exactly the real ones. Some phishing websites sends a mail to the person whose details the attacker wants to track. In this mail attacker hides his true identity and generally the attacker captures the attention of user by adding some messages and some screen captures. The innocent user's falls into this activity and provide all the important information to the attacker and became a victim of phishing. So here we introduce most secure and new method which can be used to counter the phishing attacks which is named as "An Anti-phishing framework with new interactive captcha validation scheme using visual cryptography". We provide here the method to give the user the provision to check whether the website the user wants to visit is a genuine website or a phishing website. So, by knowing these the user can securely perform his further proceedings or transactions.

Here, we combine the concept of an improved visual cryptography and new captcha scheme. Visual Cryptography (VC) is used here to divide the captcha image into two different shares and in order to get the original captcha image both of the 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 propose 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 there own advantages and disadvantages.

3. PROPOSED RESEARCH METHODOLOGY

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

The first phase in our project is the registration phase. In the registration phase, as shown in figure below we asked the user to simply enter their user-id and password for the secure website. For the most secure environment the password can be alphanumerical. After entering the user-id, password and e-mail address then the user will be seen one captcha image and is asked to enter the text shown in captcha image. Then, after entering all of the above details the user must click on the submit button to submit all of the details. As the user click on the submit button, our captcha image is divided into two shares using visual cryptography scheme we given it the names to the shares as user's share and server share. The user can download and keep it's with them. As shown in the next figure, we provide a provision to the user where the user can download it's share and the original captcha image and kept both the things with them. And, at a time of downloading user's share the another share i.e the server share is automatically sent to the any confidential or secure server and also the original captcha image along with the share as a confidential data which can be used further for verification during login phase. After the registration completed, the user can proceed to the login phase.

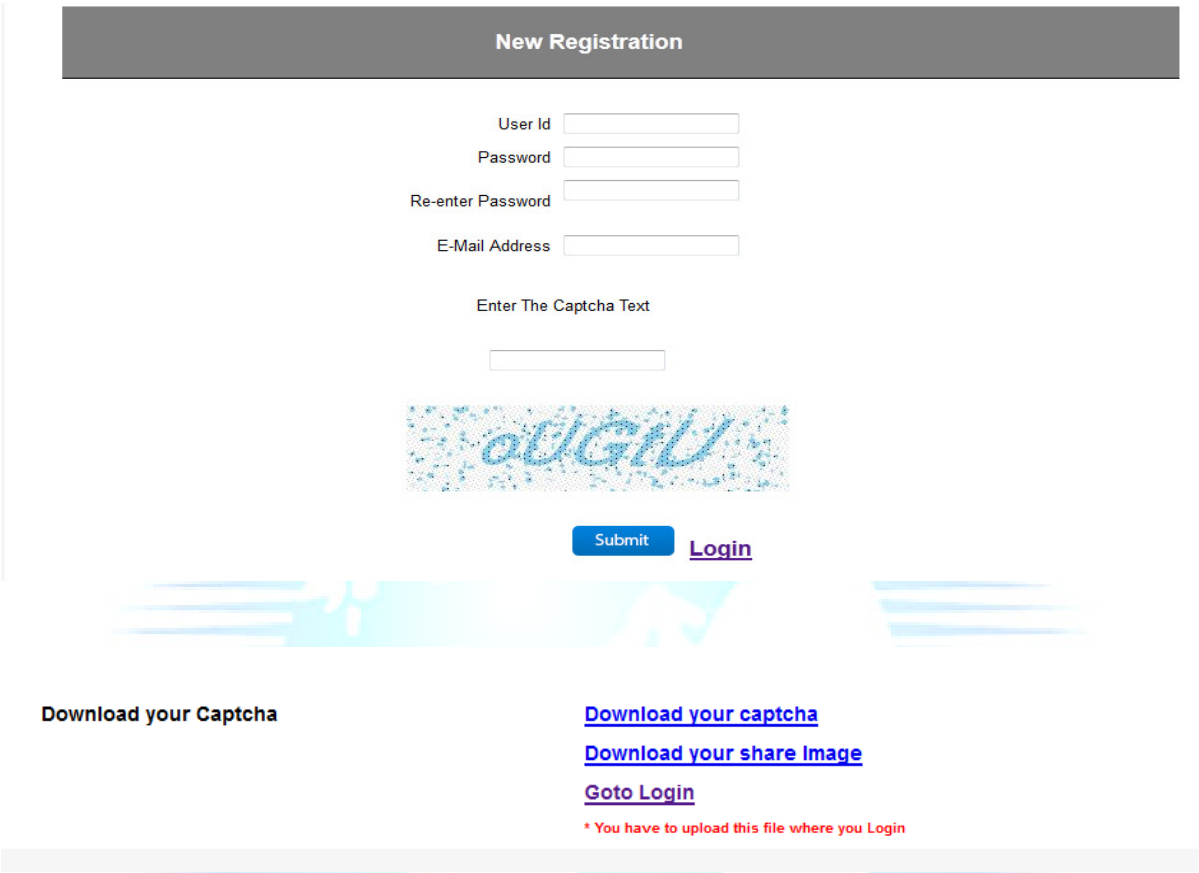

Figure 1 :- The figure showing registration phase

The second phase in our project is the login phase. In the Login phase, first the user is prompted for the username which is the user-id of the user as shown in figure below. The user is the asked to browse or to upload the user's share which is kept with the user. After uploading the user share, then the user must click on show image link as shown in figure below. As clicking on show image link, the user share is sent to the server where the user's share and the server share which is stored in a confidential server, for each user, are stacked together where one captcha image will produce. The dynamically captcha image which is generated is shown to the user. The end user here can check whether the displayed captcha image matches with the captcha image created at the time of registration. If it is same then the website the user visited is the genuine website and if it is not the same then the user knows that the website the user visited is a phishing website.

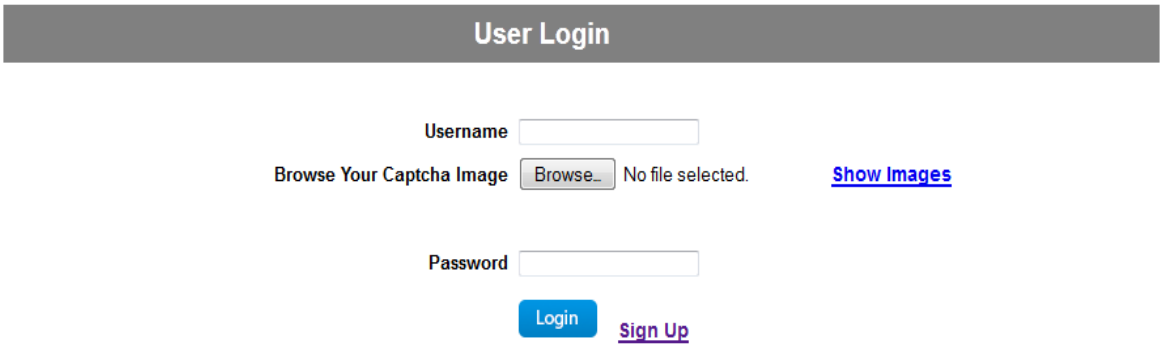

Figure 2 :- The figure showing login phase

The third and the last authentication phase of our project is the captcha solving test. In captcha solving test, the user is required to solve one captcha test. This test begins with by clicking on the captcha image as shown in figure below. After clicking on captcha image some buttons with characters are appeared below the captcha image. Here, first the user must click on the button corresponding to character-1 of the captcha image and follow the same certain sequences until all the characters in the captcha image are completed in order to successfully solve this captcha test. The idea behind this captcha test is that here we stored the per character response time. If the user is login to the website for the first time then the per character response time is stored at the server side. Next time the user is login to that website then at that time the verification process is done here i.e here the current user per character response time is compared with the previously stored per character response time which is stored at the server. If it is matches, then finally the user can successfully log in into that website and can securely perform further proceedings. If it is not matches then the user has to login again in order to access the website.

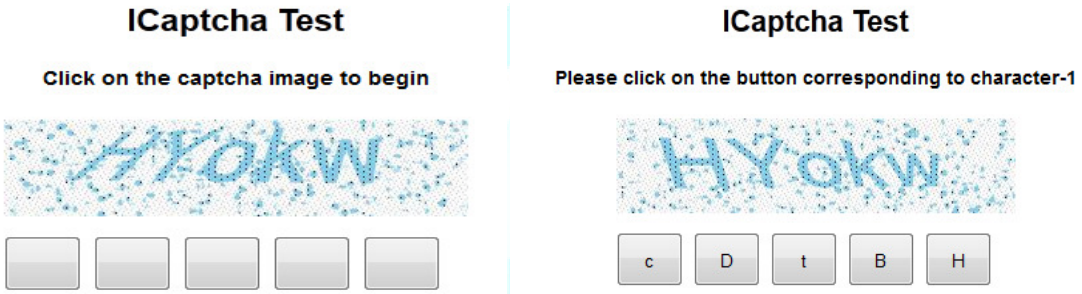

Figure 3 :- The figure showing Icaptcha test

In our proposed method, we are using here captcha solving test in order to defend against image based attacks and the third party human based attacks. Because, some 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 newly proposed systems used visual cryptographic schemes and newly designed captcha scheme in order to counter the phishing websites and the phishing attacks.

4. RESULTS

The screen below appears after user successfully completes the registration phase, login phase and after the user successfully solving or completing the captcha test.

Figure 4 :- The screen showing successful login of the user

The screen below appears if the user fails to solve the captcha test within captcha solving time or within server stored response time.

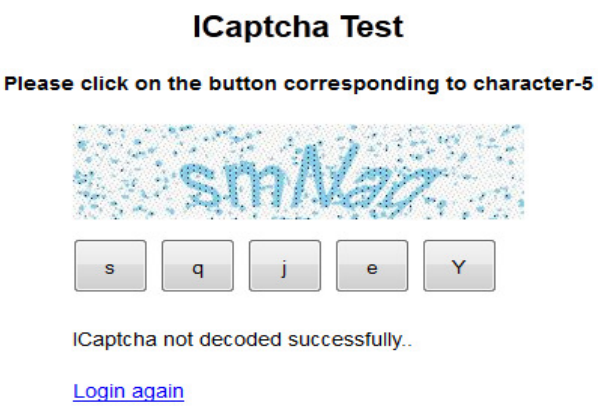

Figure 5 :- The screen showing the login failure of the user

5. CONCLUSION

In this paper, we have proposed a new method in which CAPTCHA plays an important role in our project by 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. But some researchers proposed that the current existing captcha can be attacked by using various well-known techniques such as image recognisation techniques and by using third party human solver and within submission time it can be known to the attackers. So, in our proposed method, we are using here captcha solving test in order to defend against this image based attacks and the third party human based attacks. And, in order to solve most of the problems on the internet strongly we propose to use visual cryptographic scheme and a new captcha solving scheme that can defend all the phishing attacks and can thwart all the other 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