© 2015 IJSRSET | Volume 1 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Anti-Phishing Image Captcha Validation Scheme using Visual Cryptography

R Abinaya, S Janani, P Nathiya Devi
Department of Computer Science, Dhanalakshmi College of Engineering, Chennai, Tamilnadu, India

ABSTRACT

Evolution in the world of internet has given rise to several online attacks and the most common attack is phishing. Victims are tricked into providing such information by a combination of spoofing techniques and social engineering. Phishing is an attempt by an individual or a group to acquire sensitive information such as usernames, passwords, and credit card details from unsuspecting victims. In this paper we have proposed a new approach named as "Anti-phishing image captcha validation scheme using visual cryptography" to solve the problem of phishing. Phishing websites are targeting the customers of banks and online payment services to steal sensitive data from victims. The use of images is explored to preserve the privacy of image captcha by decomposing the original image captcha into two shares that are stored in separate database servers such that the original image captcha can be revealed only when both are simultaneously available, the individual sheet images do not reveal the identity of the original image captcha. Once the original image captcha is revealed to the user it can be used as the password. Several solutions have been proposed to tackle phishing. Dynamically generating the captcha image by the system is one of the major advantages of the system. In this study, the authors shed light on the important features that distinguish phishing websites from legitimate ones and assess how good rule-based data mining classification techniques are in predicting phishing websites and which classification technique is proven to be more reliable.
Keywords: Phishing, sensitive data, image captcha, visual cryptography

I. INTRODUCTION

Online transactions are nowadays often preferred by everyone and there are various attacks present behind this. In these types of various attacks, phishing is one of the major security threat and new innovative ideas are arising in phishing for each second so preventive mechanisms should also be so effective and efficient.
Thus the security in these cases be very high and should not be easily traceable with implementation easiness. Today, most applications are only as secure as their underlying system. Since the design and technology of middleware has improved steadily, their detection is a difficult problem.

As a result, it is nearly impossible to be sure whether a computer that is connected to the internet can be considered trustworthy and all our sensitive details are secure or not. Phishing scams are also becoming a problem for online banking and e-commerce users. The question is how to handle such applications to provide a high level of security.

Phishing is a form of online procures that aims to hustle sensitive information such as online banking passwords and credit card information from users. Phishing scams have been receiving extensive press coverage because such attack has been escalating in number and sophistication.

One definition of phishing is given as “it is a criminal activity using social engineering techniques. Person attempt to fraudulently pilfer sensitive information, such as passwords and credit card details, by masquerading as a legitimate person or business in an electronic communication”. The conduct of identity theft with this acquired sensitive information has also become easier with the use of technology and identity theft can be

described as “a crime in which the swindler obtains key pieces of information such as Social Security and driver's license numbers and uses them for his or her own gain”. Phishing attacks rely upon a mix of technical deceit and social engineering practices. In the majority of cases the phisher must persuade the victim to intentionally perform a series of actions that will provide access to confidential information.

Communication channels such as email, webpage, IRC and instant messaging services are popular. In all cases the phisher takes the part of trusted source for the victim to believe. To date, the most successful phishing attacks have been initiated by email – where the phisher impersonates the sending authority, So here introduces a new method which can be used as a safe way against phishing which is named as "A novel approach against Anti-phishing using visual cryptography". As the name describes, in this approach website cross verifies its own identity by requesting for image captcha and proves that it is a legitimate website (to use bank transaction, E-commerce and online booking system etc.) before the end users and make the both the sides of the system secure as well as confidential. The concept of image processing and an improved visual cryptography is used. Image processing is a technique of processing an image input and to get the output as either improved form of the same image with the combinations of pixel and/or characteristics of the input image. Visual Cryptography (VC) is a method of embedding text into image and encrypting it into a secret image forming shares, such that stacking a sufficient number of shares reveals the secret image.

II. METHODS AND MATERIAL

A. Problem Statement

Phishing web pages are counterfeit web pages that are created by malicious people to mimic web pages of real websites. Some phishing scams use JavaScript commands in order to alter the address bar. This is done either by placing a picture of a legitimate URL over the address bar, or by closing the original bar and opening up a new one with the legitimate URL. Most of these kinds of web pages have high visual similarities to flimflam their victims. Some of these kinds of web pages look exactly like the real ones. Victims of phishing web pages may expose their bank account, password, credit card number, or other credential information to the phishing web page owners. It includes techniques such as tricking customers through email and spam messages, man in the middle attacks, installation of key loggers and screen captures. Attempts to deal with the growing number of reported phishing incidents include legislation, public awareness, and technical security measures. Phishing is a recurrent threat that keeps growing to this day. The risk grows even larger in social media such as Facebook, Twitter, Myspace etc.

Hackers commonly use these sites to attack persons using these media sites in their workplace, homes, or public in order to snag personal and security information that can affect the user. Phishing is used to portray trust in the user since the user may not be able to tell that the site being visited or program being used is not real and when this occurs is when the hacker has the chance to access the personal information such as passwords, usernames, security codes, and credit card numbers among other things.

MERITS AND DEMERITS

1. Blacklist-based technique with low false alarm probability, but it cannot detect the websites that are not in the blacklist database. Because the life cycle of phishing websites is too short and the establishment of blacklist has a long lag time, the accuracy of blacklist is not too high.
2. Heuristic-based anti-phishing technique, with a high probability of false alarm, and it is easy for the attacker to use technical means to avoid the heuristic characteristics detection.
3. Similarity assessment based technique is time-consuming. It needs too long time to calculate a pair of pages, so using this method to detect phishing websites on the client terminal is not suitable. And there is low accuracy rate for this method depends on many factors, such as the text, images, and similarity measurement.

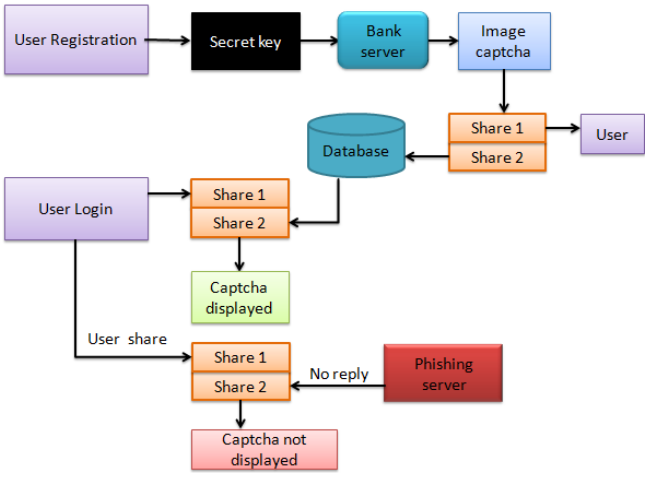

Figure 1: System Architecture

B. Proposed System

The concept of image processing and an improved visual cryptography is used. Image processing is a technique of processing an input image and to get the output as either improved form of the same image and/or characteristics of the input image. In Visual Cryptography (VC) an image is decomposed into shares and in order to reveal the original image appropriate number of shares should be combined.
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. Figure.1 denotes the shares of a white pixel and a black pixel. 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.

MERITS AND DEMERITS

1. For phishing detection and prevention, we are proposing a new methodology to detect the phishing website. Our methodology is based on the Anti-Phishing Image Captcha validation scheme using visual cryptography.
2. It prevents password and other confidential information from the phishing websites.
3. URL address on the address bar of your internet browser begins with "https"; the letter's' at the end of "https" means 'secured'.
4. Look for the padlock symbol either in the address bar or the status bar (mostly in the address bar) but not within the web page display area. Verify the security certificate by clicking on the padlock.

C. Methodology

Visual cryptography

Visual Cryptography is a special encryption technique to hide information in images in such a way that it can be decrypted by the human vision if the correct key image is used. The technique was proposed by Naor and Shamir in 1994. Visual Cryptography uses two transparent images. One image contains random pixels and the other image contains the secret information. It is impossible to retrieve the secret information from one of the images. Either transparent images or layers are required to reveal the information. The easiest way to implement Visual Cryptography is to print the two layers onto a transparent sheet.

When the random image contains truly random pixels it can be seen as a one-time pad system and will offer unbreakable encryption. In the overlay animation you can observe the two layers sliding over each other until they are correctly aligned and the hidden information appears. To try this yourself, you can copy the example layers 1 and 2, and print them onto a transparent sheet or thin paper. Always use a program that displays the black and white pixels correctly and set the printer so that all pixels are printed accurate (no diffusion or photo enhancing etc.). You can also copy and paste them on each other in a drawing program like paint and see the result immediately, but make sure to select transparent drawing and align both layers exactly over each other.

Each pixel of the images is divided into smaller blocks. There are always the same number white (transparent) and black blocks. In the table on the right we can see that a pixel, divided into four parts, can have six different states. If a pixel on layer 1 has a given state, the pixel on layer 2 may have one of two states: identical or inverted to the pixel of layer 1. If the pixel of layer 2 is identical to layer 1, the overlaid pixel will be half black and half white. Such overlaid pixel is called grey or empty. If the pixels of layer 1 and 2 are inverted or

opposite, the overlaid version will be completely black. This is an information pixel.

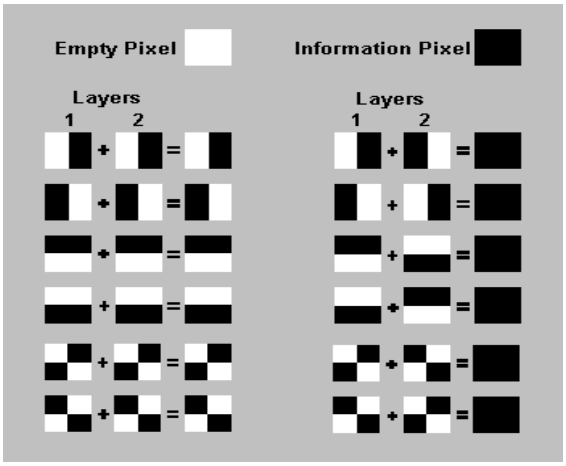

Figure 2: Separation of pixels into blocks

D. Algorithm

Linear Programming Algorithm:

This method is to achieve the best outcome in a mathematical model whose requirements are represented by linear relationships. Standard form is the usual and most intuitive form of describing a linear programming problem. It consists of the following three parts:

- A linear function to be maximized

e.g.

$$f(x_1, x_2) = c_1 x_1 + c_2 x_2$$

- Problem constraints of the following form

e.g.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &\leq b_1 \\ a_{21}x_1 + a_{22}x_2 &\leq b_2 \\ a_{31}x_1 + a_{32}x_2 &\leq b_3 \end{aligned}$$

- Non-negative variables

e.g.

$$\begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

The problem is usually expressed in matrix form, and then becomes:

$$\max\{c^{\mathrm{T}} x \mid Ax \leq b \wedge x \geq 0\}$$

Other forms, such as minimization problems, problems with constraints on alternative forms, as well as problems involving negative variables can always be rewritten into an equivalent problem in standard form.

$$G = \begin{array}{|c|cc|} \hline & m & t \\ \hline e & 3 & -1 \\ s & -2 & 1 \\ \hline \end{array}$$

The current opinion is that the efficiency of good implementations of simplex-based methods and interior point methods are similar for routine applications of linear programming. We can store the black and white pixel values by means of matrix.

III. RESULTS AND DISCUSSION

Performance Analysis

The overall performance of the proposed schemes is estimated by implementing the search system on a cloud server. The document set is built from the real data set: Reuters News stories. This dataset is a collection of 18, 821 newsgroup documents including 11, 293 train documents and 7, 528 test documents.

This product is combination of our main components, namely Image processing and visual cryptography, the web portal, web services and the JEE application. The main objective is predicting the phishing sites based on visual cryptography. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner.

The performance of the scheme is evaluated regarding the accuracy of the proposed keyword extraction method, as well as the performance of the proposed search approach.

As a result of the research tasks proposed for this project, we expect to advance the security and reliability in the following areas:

- Development of a visual cryptography methodology for specifying linear programing algorithms to improve the storage of sensitive data by imposing two step verification.
- Development of these new techniques for the implementation of security in e-banking on will naturally support the representation of high-level goals.
- The declarative language will be supported by tools for automatic transformation of specifications into target action description models.

s.no	Comparable terms	Usage	Security
1	Existing project	4.3	2.5
2	Visible captcha	4.1	3.4
3	Invisible captcha	3.5	4.3

Table 1: Performance analysis

The above given below table illustrates the security based on the methodology and techniques followed by the performance analysis graph:

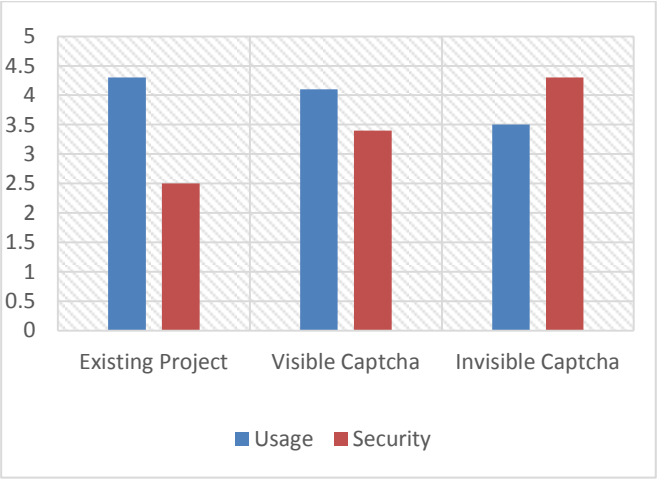

Figure 3: Performance Graph

IV. CONCLUSION AND FUTURE ENHANCEMENT

This paper mainly enlightens new anti-phishing approach to prevent damages ranges from denial of access to substantial financial loss. The flaw in the existing techniques is that the link to the website is crafted to carry out the attack, making it very difficult to spot without specialist knowledge. Phishing websites as well as human users can be easily identified using our proposed paper. This proposed methodology preserves confidential information of users with help of image shares. It verifies whether the website is a genuine/secure website or a phishing website before revealing the sensitive information.

If the website is a phishing website, then in that situation, the phishing website can't display the image captcha for that specific user who wishes to log in into the website. This is due to the fact that the image captcha is generated by the stacking of two shares, one with the user and the other with the actual database of the website. The proposed methodology is also useful to prevent the attacks of phishing websites on financial web portal, banking portal, online shopping market.

V. REFERENCES

[1] OllmannG.The Phishing Guide Understanding & Preventing Phishing Attacks, NGS Software Insight Security Research.

[2] Roberto De Prisco and Alfredo De Santis., On the Relation of Random Grid and Deterministic Visual Cryptography.

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

[4] A. Shamir, .How to Share a Secret, Communication ACM, vol. 22, 1979, pp. 612-613.

[5] G. R. Blakley, .Safeguarding Cryptographic Keys, Proceedings of AFIPS Conference, vol. 48, 1970, pp. 313-317.

[6] A. Menezes, P. Van Oorschot and S. Vanstone, .Handbook of Applied Cryptography,. CRC Press, Boca Raton, FL, 1997.

[7] Tzung-Her Chen; Chang-Sian Wu; Wei-Bin Lee,A Novel Subliminal Channel Found in Visual Cryptography and Its Application to Image Hiding.

[8] Hegde, C.; Manu, S.; DeepaShenoy, P.; Venugopal, K.R.; Patnaik, L.M. ,Secure Authentication using Image Processing and Visual Cryptography for Banking Applications.

[9] Jithi, P.V.; Nair, A.T.,Progressive visual cryptography with watermarking for meaningful shares.

[10] Kamath, M.; Parab, A.; Salyankar, A.; Dholay, S., Extended visual cryptography for color images using coding tables.

[11] Shuo-Fang Hsu; Yu-Jie Chang; Ran-Zan Wang; Yeuan-Kuen Lee; Shih-Yu Huang, Verifiable Visual Cryptography.