

Image Based Mutual Authentication to Enhance the Web Security

K. Suresh Kumar^{1*} and T. Sasikala²

¹Research Scholar, Anna University, Chennai, India, Associate Professor Saveetha Engineering College, Chennai, Tamil Nadu, India.

²Principal, SRR Engineering College, Padur, Chennai, Tamil Nadu, India.

¹sureshkumar@saveetha.ac.in, ²Sasi_madhu2k2@yahoo.co.in

Abstract

CAPTCHAs are an imperative and broadly used modern Internet technology to reduce the risk of reducing the computerized agents to programmatically use web connected resources such as online e-mail accounts, online polls, Web-based comment systems, Web-based SMS portals and so on. The designing of security algorithms utilizing captcha can be a useful security mechanism. There are several security attacks and survival of many phishing websites. In current systems, two factor authentications, One Time Password (OTP) are used to overcome these issues. In case of OTP, mobile technology is used to send the password to the client for confirmation, in some cases, mostly handheld mobiles may not be within the reach or there may be several other issues like network traffic and delay in receiving passwords. These problems decrease the capability of authentication. Here the proposed system is technology independent and behavior based authentication system, where system performs complete authentication through website using Image CAPTCHA. Behavior based authentication is done through appropriate selection and cropping of image at the accurate location. This proposed system ultimately enables us to substantiate whether the user is an authenticated user and website is an authenticated website. The performance of the system has been analyzed using shoulder surfing, brute force, dictionary, random and reply attack to prove the web security.

Keywords: Web security, mutual authentication, region selection, image CAPTCHA

1. Introduction

1.1. Overview of the Paper

Web Security has grown more essential over the years as the web has become the ultimate place for the conduct of business in today's world. There are many attacks reported worldwide that hamper web security by causing a serious threat to precious user data. One among them is the phishing attack. It is a way by which an attacker attempts to steal sensitive user information such as usernames, passwords, and other personal details by building fake websites and disguising themselves as the legitimate one.

There are various sources by which a user can be attracted towards a phishing site. Emails from unknown persons or even known persons may contain the phishing website URL that the user thinks as a legitimate one. The URL thus provided might have some extra text or data padded up compared to a legitimate URL. When these links are clicked the user ends up landing in a phishing zone that steals sensitive information. A phishing link may also appear as a search result from a web search engine. Similarly, there are a vast number of sources to lure the user into a phishing website to capture sensitive data.

Phishing scams were mainly done targeting sensitive information regarding bank accounts, credit cards of individuals, organizations, etc., People around the globe lose

* Corresponding Author

millions of dollars each year being unaware of this kind of fraudulent activities. It is important to keep the user data immune to phishing attacks with a mechanism that prompts the web users to stay safe from phishing attacks. Hence, it is necessary for business to devise an effective phishing website detection mechanism to ensure the website that the user is bound to use is a legitimate one.

CAPTCHA based approaches are effective against all these attacks. CAPTCHA is a program that can generate a rating test where humans can achieve the test, but the test cannot be passed by the current computer programs. For example in Figure 1, (Sample image of CAPTCHA), humans can read an inaccurate or distorted text, but computer programs cannot do the same.



Figure 1. Image of CAPTCHA

The applications of CAPTCHAs involves stopping comment spam in blogs, shielding website registration, preventing dictionary attacks, secure against email worms and spam. Using Image CAPTCHA as password, the user is authenticated based on examining the behavior of an individual.

1.2. Objective

The prime objective of this paper is to find whether a user is a legitimate or an illegitimate user and avert the use of phishing websites by the web user. The technique also grants confirmation for the web users for assuring websites that was offered to them. It guarantees to keeps the system, technology independent. This technique provides security by using a novel method that was based on a cropping image using symbols that were being implemented during registration to avert phishing websites and to allow the legitimate user to get logged in.

The ICS (Image CAPTCHA Cropping using Symbols) algorithm also provides authentication for the authorized users. The ICS algorithm checks the honesty of the visited website, and it also checks the trustworthiness of the user. It allows only the approved users to enter using the condition, an image cropping process to the user by following the ICS algorithm. The ICS algorithm records and cross checks the behavior of the user for assuring the integrity of the user and it also sends the data in a secured way using RSA and AES algorithm.

2. Literature Review

Literature presents several techniques for Image-based CAPTCHA technology for enhancing web security. Here we review some of the techniques presented in the literature in the recent times. Anthony Y. Fu, *et. al.*, [1] have presented their work for 'Detecting Phishing Web pages with Visual Similarity Assessment'. It's a useful approach for detecting phishing webpages which employ the earth mover's distance to calculate the visual similarity of web pages. EMD is used to calculate the signature distance of images in the web page. If EMD based visual similarity of the suspicious web page exceeds the threshold value then it is a phishing web page. Barry Leiba [2] has presented his work on 'OAuth Web Authorization Protocol', the OAuth protocol provides a secure way to authenticate services in the network without exposing the personal identities of the users associated with their account. The process was done by providing tokens to the users which limit the area of access or the time of access and other related features for using the desired service. Chun-Ming Leung [3] has proposed his work on "Visual Security Is Feeble for Antiphishing," After analyzing the current CAPTCHA technique; the system suggests improved security techniques using properties of the CAPTCHA along with the restrictions of one-time password.

Danesh Irani, *et. al.*, [4] have presented their work on "Modeling Unintended Personal Information Leakage from Multiple Online Social Network" An empirical study is made to evaluate identification attack and password recovery attack. The result shows that security risk is higher for multiple account users than single account users. Thus quantifying the amount of information a user's social footprint is made. David Geer [5] has proposed a new authentication scheme called Passmark security. The email from the legitimate websites contains the phrase along with the image provided by the user during registration. When the user enters the details along with the image during login if the suspect page is a legitimate page it should send the same phrase to the user received through email. Eric Grosse *et. al.*, [6] presented their paper on "Two step Verification" As the ordinary text based passwords are not secure. To ensure security an additional authentication called two step verification that uses text based passwords along with the data sent to the user device has been used for authentication Hung-Min Sun *et. al.*, [7] in their paper developed a User authentication protocol called Opass which is designed to prevent password stealing and reuse attacks that leverage user's cell phone and SMS. Using Opass users only need to remember a long term password for login on all websites. If the cell phone has been missed, it can be retrieved by reusable sim cards and long term password.

Kuan-Ta Chen *et. al.*, [8] have proposed their work to detect the phishing websites using Discriminative Key point Features. The snapshot of the suspect web page is compared with the snapshot of the legitimate webpage in the database. Then feature matching is done using key point feature calculator and extractor. Based on the threshold values of the page similarity it is decided whether the suspect webpage is a phishing website or not. LiuWenyin, *et. al.*, [9] have presented their paper on "Antiphishing through Phishing Target Discovery" The snapshot of the suspect web page has been compared with the snapshot of the legitimate webpage in the database. Feature matching is done using key point feature calculator and extractor. Based on the threshold values of the page similarity decision is made whether the suspect webpage is a phishing website or not. Ritendra Datta, *et. al.*, [10] has proposed their work to show the difference between human and computer using CAPTCHA. The recognizing ability of human and machine is differentiated using an Automated Turing test. Human recognizing ability is measured using an extensive user study and the ability of recognition by the machine through content based image retrieval (CBIR) technique. The CAPTCHAs are easily recognized by the humans only and not by computer programs.

Saikat Chakrabarti, *et. al.*, [11] has presented their paper on preventing dictionary attacks. Acceptable password protocol geared to prevent dictionary attack was proposed. The attack was defeated by using CAPTCHA that asks the users to solve visual Recognition pattern making life difficult for the machine. Sonia Chiasson, *et. al.*, [12] have proposed a Password authentication system. To avoid the problems in knowledge based authentication (text based passwords) a password authentication system named persuasive cued click points was developed. The system provides strong passwords generated by the system in the form of click images using viewport based on tolerance square. The method is harder to guess and easy to remember thus maintaining remembrance, usability, and security. Wei-Bin Lee, *et. al.*, [13] has proposed their work based on Anti Phishing scheme, where the mutual authentication was done between the user and the server. As we need a sensitive key table for achieving this, here, the system is developed without using the key table. This technique also helps them to overcome several attacks.

3. Proposed System Methodology

The proposed system focuses on identifying phishing websites to the users and to provide secure authentication by allowing access to authorized persons. This has been done by image CAPTCHA via Image CAPTCHA Cropping using Symbols (ICS) algorithm as shown in Figure 2. During Registration after entering all the details, the user has to select three images from 80 random images generated by the server and needs to perform the cropping process in canvas.

The cropping should be done using the symbols, each image has been provided with five symbols, and the user can choose any one of those symbols to place it anywhere on an image for the cropping process. The behavior of the user in the selection of images, symbols, and placing of symbols were recorded and the cropped pieces are then stored along with the registration details in the database of the server. During login after entering the username if the client receives anyone of the images selected during registration along with nine random images, the user can identify that the suspicious webpage is a legitimate website or it is a phishing website. If the suspicious website is a legitimate one, then the user has to perform the cropping process in the same way done during registration. The behavior of the user during the login process has been recorded and the cropped images are encrypted using RSA algorithm. Then the encrypted details are sent to the server.

The server has to use the ICS algorithm to measure the similarities between the cropped pieces of the image during login and the result stored in the database during registration based on the position of a cropped image by measuring the deviation of the cropped image on the original image's center point and radius. After measuring the similarities, the ICS algorithm checks the overall accuracy of the cropped pieces as well as the individual accuracy of the cropped pieces.

Based on the selection of images, symbols and the position of cropped pieces, the error rate is incremented. If the accuracy is greater than 99% it provides password link for the users to enter text password else if the accuracy is greater than 80% in the first attempt; it provides second chance to crop. If the accuracy is greater than 90% in the second attempt, it provides third chance to crop. If the accuracy of the cropped pieces is greater than 99% in the third attempt, the user is allowed to enter his password in the password link. If the accuracy is less than 80% during the first attempt or less than 90% during the second attempt, it blocks the access to the user. The user can also change the random image after his login if the user wishes to do. If the cropped pieces accuracy is the similar percentage for three times consecutively of the user login an alert will be sent to the user to provide more accurate cropping.

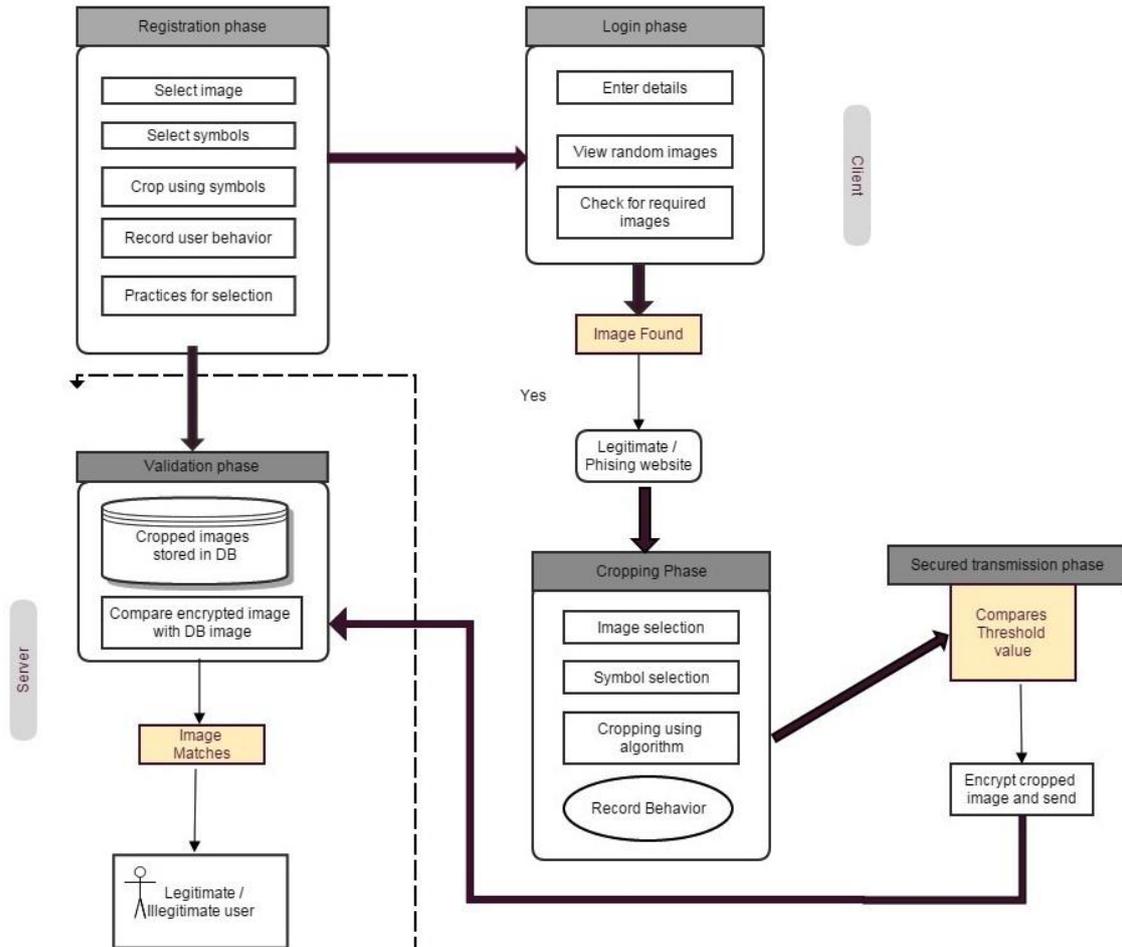


Figure 2. Block Diagram of Proposed System

3.1. Image Selection During Registration

The client fills all the personal details and after entering the personal details, it will be displayed with 80 images. Out of the 80 images, the client selects three images.

$$Nima_d = ima1_d, ima2_d, ima3_d,$$

N= Number of selected images

After selecting the images, the client selects the symbols for performing the procedure of cropping

$$Nimas_d = imas1_d, imas2_d, imas3_d.$$

The process of cropping is defined as,

$$Crp(1...N)_d == Crp(3)_d == crp1_d, crp2_d, crp3_d$$

The server records all the behaviour of the user based on their selection Implementation on and cropping of images. Practice sessions are also given to them to gain practice on the process of cropping.

3.2. Login and Detecting Phishing Websites

The user id is defined by UID. the user is allowed to login only if user is a registered user.

$$\text{Primary login} = \begin{cases} \text{Next level, if } (u_i = UI_D) \\ \text{Not Remain, if } (u_i \neq UI_D) \end{cases}$$

Step 1: Show three of the right image, along with 30 random images

Step 2: Check the right images $ima1_d, ima2_d, ima3_d$ for the received username (UID).

$$\text{Login} = \begin{cases} \text{Genuine website, if images} = Nima_d \\ \text{Phishing website, if images} \neq Nima_d \end{cases}$$

Step 3: User chooses similar images ($ima1_b, ima2_b, ima3_b$) selected during registration process Images chosen during login, $Nim_b = (ima1_b, ima2_b, ima3_b)$

3.3. Image Cropping and Behavior Checking

The client performs the cropping process on the selected images-**Nil**. The image is cropped using symbols, $ims1_i, ims2_i, ims3_i$.

ICS algorithm measures the similarities between the images cropped during login and the images chosen during registration and behavior check is done.

$$\text{Crp}(1\dots N)_i = \text{Crp}(3)_i = \text{crp1}_i, \text{crp2}_i, \text{crp3}_i$$

$$\text{Behaviour Check} = \begin{cases} \text{Allow, If, } \mathbf{Nim}_i = \mathbf{Nim}_d, \\ \mathbf{Nims}_i = \mathbf{Nims}_d \\ \mathbf{Cr}(3)_i = \mathbf{Cr}(3)_d \\ \text{Does not allow, Else} \end{cases}$$

Maximum of three attempts given for the selection and cropping of the image. In all the attempts, the same cropping process is performed, and behaviour was checked. Only if the Error rate is less than 4%, the client will be allowed to login.

The cropping and behaviour checking algorithm works as follows:

Step 1: Start the cropping process
Step 2: Select a symbol for each image
Step 3: If cropped image is within radius
 The error rate is incremented to 0.5%
 else if cropped image overlaps the outline
 The error rate is incremented to 0.75%
 else if cropped image is outside the radius
 The error rate is incremented to 1%
 else
 Quit
Step 4: The cropped images are encrypted and sent to DB to check the overall accuracy

3.4. Secured Transmission

After performing cropping process during login, it submits the images, symbols and cropped image in encrypted form. The mechanism of Public and private key generation for encryption is done using RSA algorithm. The Rivest-Shamir-Adleman (RSA) algorithm is an encryption algorithm with two keys involved in it, one as a public key and the other as the private key. To encrypt and decrypt messages, Advance Encryption

Standard algorithm has been used. The Advanced Encryption Standard (AES) works as follows:

- Data is processed as four columns of 4 bytes.
- Depending on the key size- 128, 192, 256, the number of rounds is fixed as 10, 12 and 14.
- The entire data block is accessed on every round.

3.5. Validation and Image Authentications

The detailed data of cropped images during initial login by the user were compared with the cropped images details stored in the database.

Step 1: For every wrong selection of image, the error rate increases to 10%, defined by Δi

Step 2: For every wrong selection of symbol, the error rate increases to 5%, defined by Δs

Step 3: For every discrepancy in cropping, the error rate increases up to 1%, defined by Δc

Step 4: Total Error Rate (Δe) = $\Delta i + \Delta s + \Delta c$

Step 5: if $\Delta e > 4\%$, given with second attempts.

First attempt

The error rate of selecting the image, selecting the symbols and placing the symbols should be less than 16%, and then it goes to the second attempt.

Second attempt

The error rate of selecting the image, selecting the symbols and placing the symbols should be less than 6%, and then it goes to the third attempt.

Third attempt

The error rate of selecting the image, selecting the symbols and placing the symbols should be less than 0.5%, and then the client is allowed to submit the cropped image. If not, the user is an intruder.

Finally, the overall accuracy of the cropped images is checked. If the cropped images do not match, the user will be declared as an illegitimate user. If the cropped images match, the particular user is a legitimate user.

4. Results and Discussion

This section discusses the results obtained by this proposed technique. Some analyses on various attacks were also described here.

All the web applications can only be accessed by registering the users in the particular application. Initially, two options will be given to the user, either to register or to login. If the user has already registered, he has to directly login using his username and password. If not, the particular user has to register himself first and then, he has to proceed to login.

Normally all the login forms will only hold username and password, but our proposed technique allows the user to login only if they enter/ select their correct username, password, image, symbols and cropping of that particular image.

4.1. User Interface Design

The user will be displayed with 80 random images; the user has to select two or three images from the 80 random images as shown in Figure 4.1, a & Figure 4.1.b.

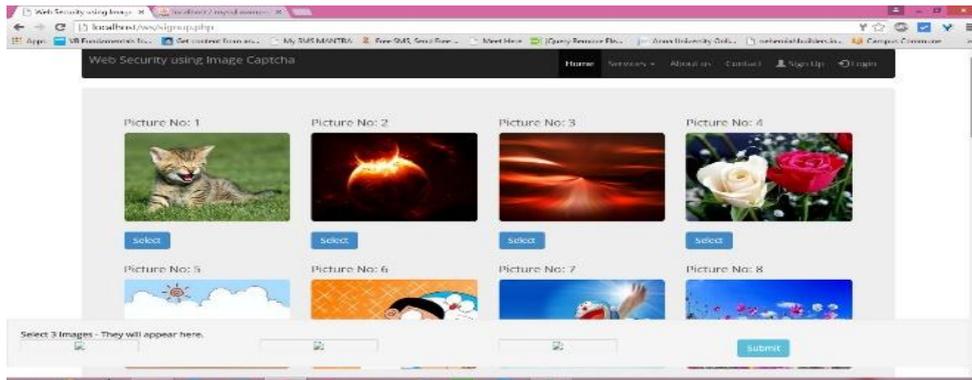


Figure 4.1.a. Random Display of 80 Images



Figure 4.1.b. Displays Selected Three Images

Followed by selecting the images the user has to involve in the process of selecting symbols for all the images selected. The user has to choose one symbol for each image as shown in Figure 4.1.c.

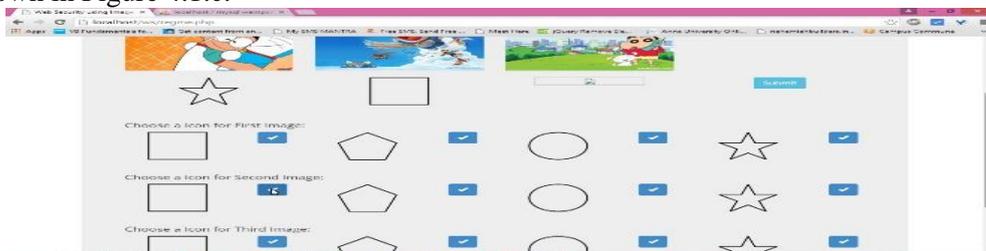


Figure 4.1 c. Displays Selected Symbols

After the selection of symbols, the user has to place each of those symbols on the corresponding images to perform cropping as shown in Figure 4.1.d.



Figure 4.1 d. Placing Symbol on the Images for Cropping

The user is given an option to go with the three practice sessions to get practice of cropping the images as shown in Figure 4.1.e. All the behavior of the users in selecting the images, symbols and placing the symbols over the images are recorded.

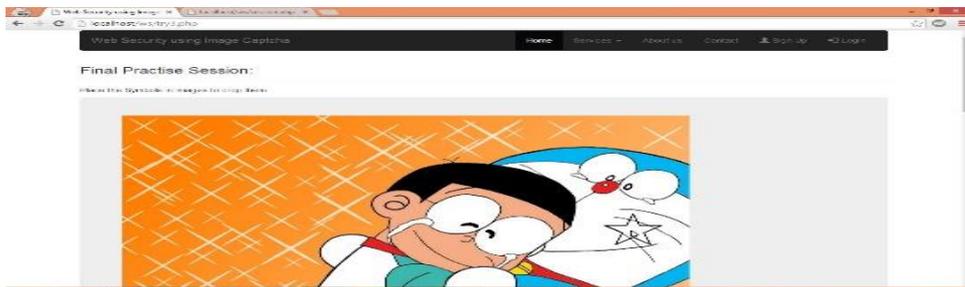


Figure 4.1 e. Practice Sessions to Crop Images

After the completion of the registration process, the user will be allowed to login, the user is given with three login attempts. During login, the user will be displayed with 30 random images, containing the images which have been selected during registration by the same user as shown in Figure 4.1.f. If those selected images during registration were not found, the user detects this as phishing website. If not the user continues with the process of login by selection of images, symbols, and cropping.

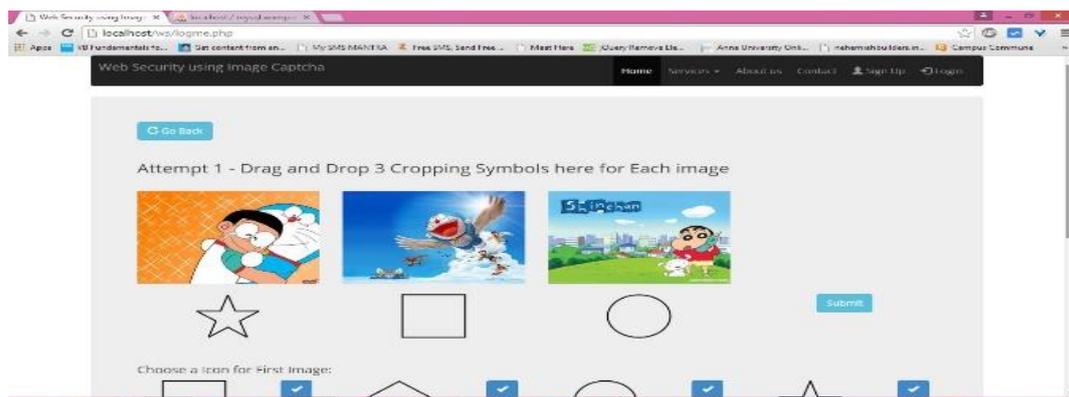


Figure 4.1.f. Selection of Images and Symbols

The behavior of the user in selecting the images, symbols and cropping the images would be compared with the data that we get during the registration process.

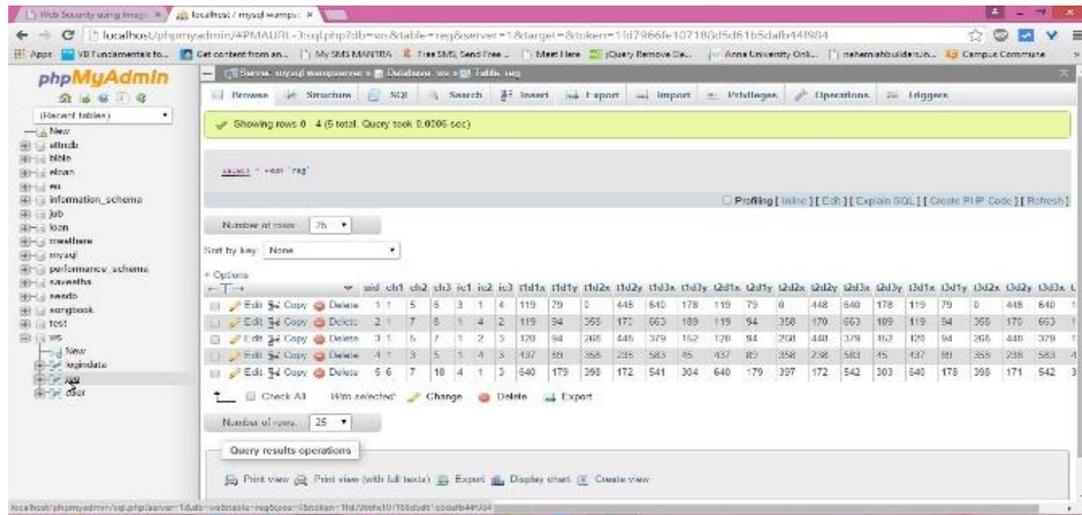


Figure 4.1.g. User Registration Details

The error rate would be incremented for every wrong selection or cropping of the image. Based on this error rate, the user either will be allowed to enter into the web application in the first, or second or third attempt, or will be asked to login again. The times taken by the user in each of the attempts are also being noted to analyze the performance of the user as shown in Figure 4.1.g & 4.1.h.

User Login Details

ID	Time	#1 Er-Rate	#1 Time	#2 Er-Rate	#2 Time	#3 Er-Rate	#3 Time	Status	Done In
1	05-5-2015 7:35:49	0.5%	48	%	0	%	0	Success	1
1	05-5-2015 7:37:14	2%	45	0.5%	49	%	0	Success	2
1	05-5-2015 7:39:21	7%	33	0.5%	46	%	0	Success	2
2	05-5-2015 7:43:14	14%	7	11%	61	%	0	Failed	0
2	05-5-2015 7:47:03	2%	43	0.5%	44	%	0	Success	2
2	05-5-2015 7:49:17	2%	368	%	0	%	0	Started	0
2	05-5-2015 7:56:05	1%	32	1%	37	0.5%	38	Success	3
3	06-6-2015 2:49:50	2%	43	1%	118	0.5%	223	Success	3
4	06-6-2015 6:22:50	0.5%	129	%	0	%	0	Success	1
5	18-18-2015 6:24:16	8%	279	%	0	%	0	Success	1
6	09-9-2015 2:10:45	3%	1822	%	0	%	0	Success	1
6	09-9-2015 2:10:45	%	-1425863445	%	0	%	0	toStart	0

Figure 4.1.h. Record of Error Rate during Training Session

4.2. Performance Analysis

We have done performance analysis with different parameters like time, success rate and attack with our proposed method. We have involved about 197 persons with different age groups and different backgrounds. Out of 197 persons 55 % are students (Category A) in the age group of 18 to 23 and the 45 % of persons are in the age of above 25(Category B).

Each user is given a practice with our system in order to get familiar with the process and were tested at various time intervals (1st Week, after One Month, 3 Months). We found that 95% of the users belonging to category A succeeded in the first attempt and 92% from the category B after the 3rd month as shown in Figure 4.2.a.

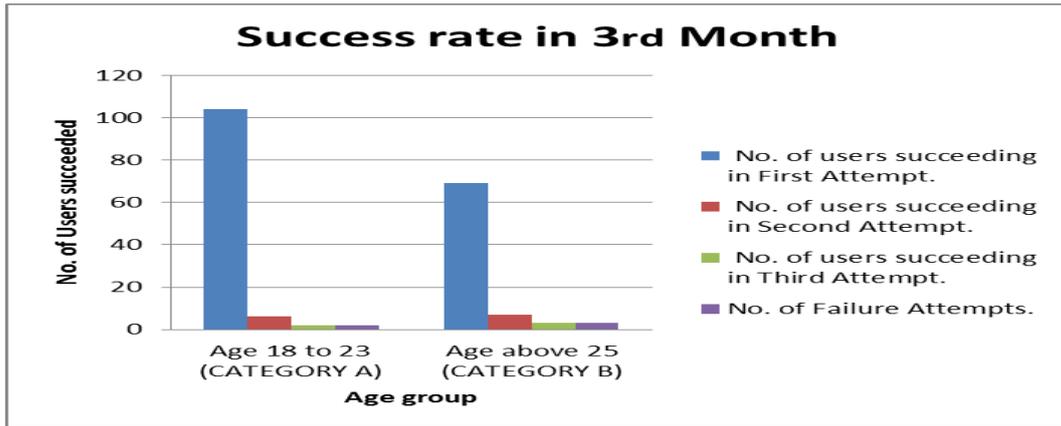


Figure 4.2.a. Success Rate of Category A & B

- Each user is requested to select three images and its corresponding symbols.
- The periodic test has been done on all the users in order to find out the success rate.

We also recorded the time taken to complete the successful attempts. At most 97% of the category A users and 94% of the category B users completed their successful attempts with the same stipulated time. The below graphs show the results.

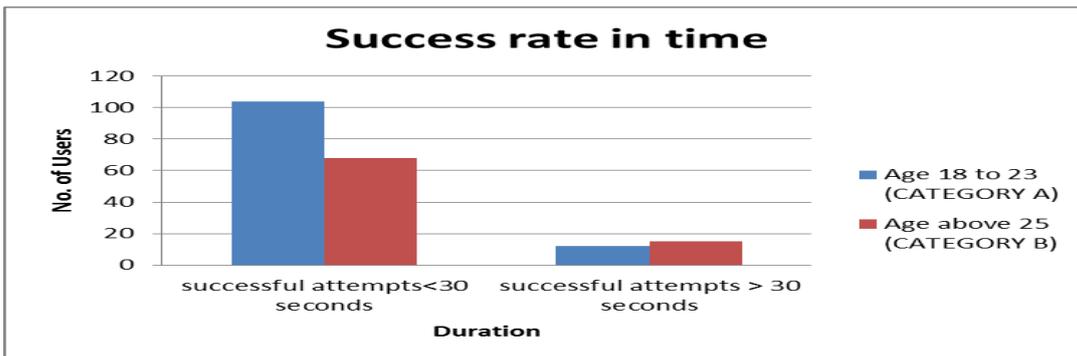


Figure 4.2.b. Success Time of Category A & B

4.3. Attack Analysis

Comparing with all other authentication mechanism, our proposed system stands against many of the attacks such as Data logger, Brute force attack, Cross-site Scripting, SQL Injection, Denial of Service, *etc.*, Our proposed system can work by overcoming all these attacks. The performance of our proposed system and other authentication mechanisms are put forth with the help of the graph as shown in Figure 4.3.

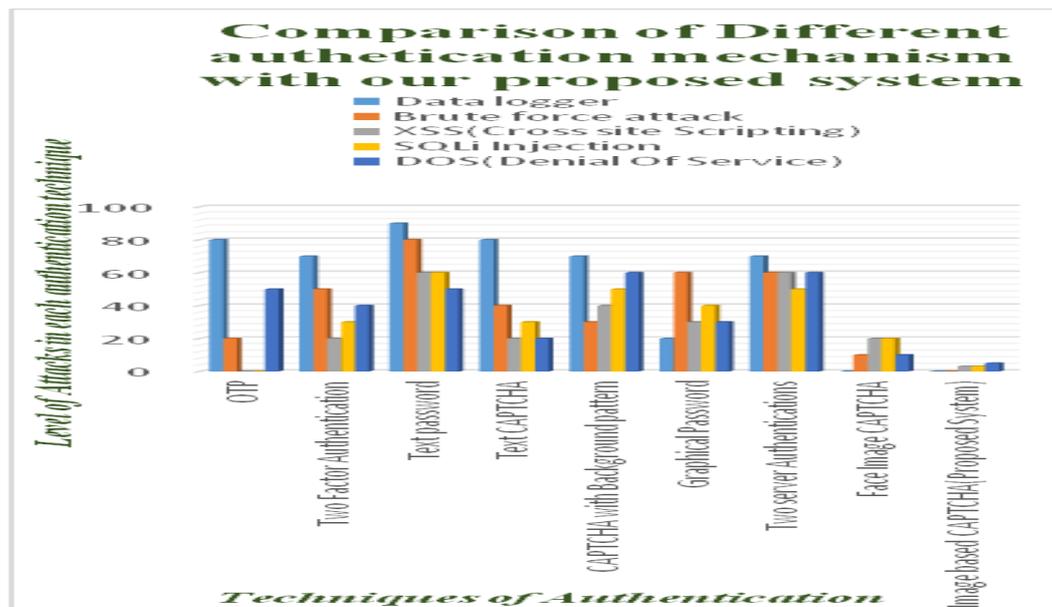


Figure 4.3. Comparison of Authentication Mechanisms

5. Conclusion

The proposed system helps the user in identifying if the user is a legitimate user or an intruder and it ensures the visited website is a phishing website or a legitimate website by using an effective image CAPTCHA system. The user is entitled to image cropping along with submitting user details. The ICS algorithm restricts user selection during cropping by giving symbols for cropping an image. The function of the ICS algorithm can be extended to provide authentication to the registered users along with phishing website detection. The algorithm would analyze the performance of the user in cropping during login and provides secured communication. The user is authenticated only after successful completion of cropping or else the user is not authenticated. The result of the proposed method indicates that it is a user friendly and secured mechanism to authenticate the user and also works as a challenge response system so that the server will also authenticate itself as a legitimate system.

Acknowledgements

The authors would like to express their sincere thanks to the management and students for their invaluable support and encouragement.

References

- [1] A. Y. Fu, L. Wenyan and X. Deng, "Detecting Phishing Webpages with Visual Similarity Assessment Based On Earth Mover's Distance", IEEE Transactions on Dependable and Secure Computing, vol. 3, (2006), pp. 301-311.
- [2] B. Leiba, "OAuth Web Authorization Protocol", Internet Computing, vol. 16, (2012), pp. 74-77.
- [3] C. M. Leung, (2009), "Visual Security Is Feeble For Antiphishing", Proceedings of the 3rd International Conference on Anti-counterfeiting and Identification in Communication, (2009), pp. 118-123.
- [4] D. Irani, S. Webb and C. Pu, "Modeling Unintended Personal Information Leakage from Multiple Online Social Network", IEEE Internet Computing, (2011), pp.13-19.
- [5] D. Geer, "Security Technology Go Phishing", IEEE, vol. 38, no. 6, (2005), pp. 18-21.
- [6] E. Grosse and M. Upadhyay, "Authentication at Scale", IEEE Security and Privacy, vol. 11, (2013), pp. 15-22.
- [7] H. M. Sun, Y. H. Chen and Y. H. Lin, "Opass: A User Authentication Protocol Resistant To Password Stealing and Password Reuse Attacks", IEEE Information Forensics and Security, vol. 7, (2012), pp. 651-663.

- [8] K. T. Chen, C. R. Huang and C. S. Chen, "Fighting Phishing with Discriminative Key point Features", IEEE Internet Computing, vol. 13, (2009), pp. 56-63.
- [9] L. Wenyin, G. Liu, B. Qiu and X. Quan, "Antiphishing through Phishing Target Discovery", IEEE Internet Computing, vol. 16, (2012), pp. 52-61.
- [10] R. Datta, J. Li and J. Z. Wang, "Exploiting the Human Machine Gap in Image Recognition for Designing CAPTCHA", IEEE transactions on information forensics and security, vol. 4, (2009), pp. 504-518.
- [11] S. Chakrabarti and M. Singhal, "Password-Based Authentication Preventing Dictionary Attacks", IEEE, vol. 40, (2007), pp. 68-74.
- [12] S. Chiasson, E. Robert, A. Forget, R. Biddle and P. C. V. Oorschot, "Persuasive Cued Click-Points: Design, Implementation, and Evaluation of a Knowledge-Based Authentication Mechanism", IEEE Transactions on Dependable and Secure Computing, vol. 9, (2012), pp. 222-235.
- [13] W. B. Lee, H. B. Chen, S. S. Chang and C. C. Yang, "An Anti-Phishing User Authentication Scheme without Using a Sensitive Key Table", Proceedings of the International Conference on Intelligent Information Hiding and Multimedia Signal Processing, (2011), pp. 141-1 .

Authors



K. Suresh Kumar, (Krishnamoorthy Suresh Kumar) obtained his Bachelor's degree in Computer Science from Bharatiyar University. Then he obtained his Master's degree in Madurai Kamaraj University and he also completed his Master of Technology in Information Technology from Sathyabama University. Currently he is pursuing PhD in Computer Science majoring in Web Security from Anna University, Chennai, Tamil Nadu, India. He is the Life Member of CSI and ISTE. Currently, he is an Associate Professor in Department of Information Technology, at Saveetha Engineering College Chennai. His specializations include Web Security, Networking and Mobile Computing. His current research interests are Web Security, Authentication and E-commerce.



T. Sasikala, She received her B.E.CSE, M.E.CSE degree, and the Ph.D. degree from the Sathyabama University, Chennai, Tamil Nadu, India. She has 20 years of experience in teaching. Currently she is the Principal of SRR Engineering College, Padur, Chennai. Her research interests are in networks, wireless sensor networks, and heterogeneous wireless networks. Dr. T. Sasikala is a Life Member of CSI. She has published 45 papers in various conferences, including the IEEE International Conference and journals. Currently 16 research scholars are doing research under her guidance.

