International Innovative Research Journal of Engineering and Technology
ISSN NO: 2456-1983

International Conference on Emerging Innovation in Engineering and Technology

ICEIET-2017

A Survey On Occluded Image Based Captcha For Online Authentication

M. Tamizh[1],Dr. M. Ezhilarasan[2]
[1]Student, [2]Professor, Information Technology[1,2]Pondicherry Engineering College,
[1]tamizhmurugan23@gmail.com

Abstract

CAPTCHA is a security mechanism that distinguishes human from malicious computer. Text based Captcha ask the user to identify distorted text. It brings complexity for the user to recognize the text. In order to overcome such difficulty Image based Captcha is developed. Image based Captcha has another type of Captcha such as Confident Captcha that asks the users to click on all images that gives a specific type of symbols such as cats or birds. Such kind of image based Captcha faces difficulty, since it could be easily identified by the unauthenticated user. The proposed system focus on providing authentication by using Occluded Image based Captcha. Occlusion is applied on images. Occlusion used is affine transformation. Affine transformation is used for providing complexity to user. Transformations such as Rotation, Translation, Scaling and Shearing will be applied to the images. Rotation is used for rotating the image to certain angle. Translation will translate the shape of an image. Shearing slants the shape of an image. Scaling will changes the size of an image. The user has to select the image during enrollment and the details will be stored into database. During login the user has to select the shuffled image that is selected in the enrollment. The selected image will be compared with database. If it matches, the user is authenticated. If the match is not detected, the user will not be authenticated. To bring complexity in authentication OTP is generated and sent to registered ID.

Keywords- Affine transformation, CAPTCHA, Image Captcha, Occlusion, OTP

I. INTRODUCTION

Web Security is the process of preventing and detecting unauthorized access to computer. It involves the process of safeguarding computers against intruders from using computer resources. Computer security is a branch of Information Security and is often used for securing the computer access. It encompasses several security measures such as software programs like anti-virus suites, firewalls, and user dependant measures such as activating deactivating certain software features like Java scripts, ActiveX in using the computer and the network resources or the Internet.

CAPTCHA technology has an experiment called the Turing Test. Alan Turing, the father of modern computing, proposed the test as a way to examine whether or not machines can think like humans. The classic test is a game of imitation. The goal of Captcha test is to create a test to distinguish that humans can pass but machines cannot pass. It is also important that the CAPTCHA application is able to present different CAPTCHAs to different users. If a visual CAPTCHA presented a static image that was the same for every user, it would not take long before a spammer spotted the form, deciphered the letters, and programmed an application to type in the correct answer automatically [1]. Most of CAPTCHAs does not rely on a visual test. Computers lack the sophistication that human beings have come to processing visual data. By looking at an image users can pick out patterns more easily than a computer. The human mind will perceive patterns easily even when none exist. But not all CAPTCHAs rely on visual patterns. In fact, it is important to have an alternative type to a visual CAPTCHA. One alternative technique to a visual test is an audible one. An audio CAPTCHA usually presents the user with a series of spoken letters or numbers. It is not unusual for the program to distort the speaker's voice, and it is also common for the program to include background noise in the recording. This helps thwart voice recognition programs. Another option is to create a CAPTCHA that asks the reader to interpret a short passage of text. A contextual CAPTCHA quizzes the reader and tests comprehension skills. While computer programs can pick out key words in text passages,

they aren't very good at understanding what those words actually mean.

A CAPTCHA has various forms like text based or image based CAPTCHA. The bot operation is similar to that of the reverse "TURING TEST", where the program will be acting like a judge and the other person will be acting like a user [7]. If the user fails the test then he/she is considered to be a machine otherwise the user is considered to be an authenticated user or a human being. CAPTCHA is a system that acts as a tool to check web Bots from exploiting online services on the internet including free email providers, wikis, and blogs. It is a HIP system that is widely used for securing the internet based applications. It is also referred to as a challenge response test which gives a challenge to the users, when the user gives correct answer then the user is considered to as a human otherwise a web bot. CAPTCHA is an authentication process that is based on challenge-response test. CAPTCHA provides a mechanism that protects the users from spam and password decryption by a simple test. In this test a user will see either an distorted image or a distorted text. The user has to enter the pattern exactly as shown to them if the CAPTCHA is based on text. If the CAPTCHA is based on image the user has to enter the correct name of the image which correctly symbolizes the image. CAPTCHA is most widely used in authentication process. Various web services like Yahoo, Google, and Bing use CAPTCHA to differentiate between an authenticated user and a malicious program. CAPTCHAs are also being used in the sites that provide access to sensitive data, such as credit card accounts and banks.

II. RELATED WORK

CAPTCHA is the test for distinguishing humans and computers are apart. The CAPTCHA can be classified into following categories:

a) Text Captcha
b) Audio Captcha
c) Video Captcha
d) Image Captcha
e) Puzzle Captcha

a) Text Captcha- It is simple and very effective method to implement. In Text based captcha the Number of characters and digits are very small so the problem occurs for user to identify the correct characters and digits. Text Captcha is easy to identify the character as well as digits through Optical Character Recognition (OCR).
b) Audio Captcha- It is based on the sound based system. It is used for blind people inorder to prove that they are not robot. It is being developed for visually disabled users. It usually contains downloadable audio-clips [7]. In this type of CAPTCHA, first the user listens and after that submits the spoken word.
c) Video Captcha- Video CAPTCHA is a newer and less commonly seen CAPTCHA system. In video-based CAPTCHAs, three words are being provided to the user that describes a video. The user's tag must match to a set of automatically generated tags then the test is said to be passed.
d) Image Captcha- Graphics-based CAPTCHAs are challenge-tests in which the users have to guess those images that have some similarity [1]. The advantage of image based CAPTCHA is that pattern recognition is hard AI problem and therefore it is difficult to break the test using pattern recognition technique.
e) Puzzle Captcha- Usually in puzzle based CAPTCHA a given picture is divided to chunks [7]. A user is supposed to combine these chunks so as to form the complete picture same as the original one.

III. LITERATURE SURVEY

The following are the Research paper I have studied, that is given below

Guido Schryen, Gerit Wagner , Alexander Schlegel [13], In their Research paper they have mentioned two face recognition Captcha and security level of both. The level of security is provided by text-based CAPTCHAs. At the same time, techniques for distorting and obscuring the text, which is used to maintain the level of security, make text-based CAPTCHAs difficult to solve for humans, and thereby further degrade usability. The needs for developing alternative types of CAPTCHAs improve both the current security and the usability. The two novel face recognition Captcha is based on gender and age classification. The accuracy of gender recognition algorithms strongly depends on the quality of images. Face recognition Captcha can be strengthened by increasing the number of images, applying image distortion techniques and/or selecting images that are difficult to classify.

Bin B. Zhu, Jeff Yan, Guanbo Bao, Maowei Yang, and Ning Xu [4], in their Research paper they have used a new security primitive based on hard AI problems, namely, a novel family of graphical password systems built on top of Captcha technology, which call Captcha as graphical passwords (CaRP). CaRP is both a Captcha and a graphical password scheme. CaRP addresses a number of security problems altogether, such as online guessing attacks, relay attacks, and, if combined with dual-view technologies, shoulder-surfing attacks. Notably, a CaRP password can be found only probabilistically by automatic online guessing attacks even if the password is in the search set. CaRP also offers a novel approach to address the well-known image hotspot problem in popular graphical password systems, such as PassPoints that often leads to weak password choices. CaRP is not a

panacea, but it offers reasonable security and usability. The notion of CaRP introduces a new family of graphical passwords that adopts a new approach to counter on-line guess attacks: a replacement CaRP image, that is additionally a Captcha challenge, is used for every login arrange to create trials of an internet guess attack computationally freelance of every different. A password of CaRP is found solely probabilistically by automatic online guess attacks together with brute-force attacks, a desired security property that different graphical word schemes lack. Hotspots in CaRP pictures will now not be exploited to mount automatic on-line guess attacks, AN inherent vulnerability in many graphical word systems.

Hung-Min Sun, Shiuan-Tung Chen, Jyh-Haw Yeh and Chia-Yun Cheng [14], in this paper Attackers can observe directly or use external recording devices to collect users credentials. To overcome this problem, they proposed a novel authentication system PassMatrix, based on graphical passwords to resist shoulder surfing attacks. With a one-time valid login indicator and circulative horizontal and vertical bars covering the entire scope of pass-images, PassMatrix offers no hint for attackers to figure out or narrow down the password even they conduct multiple camera-based attacks. To overcome the security weakness of the traditional PIN method, the easiness of obtaining passwords by observers in public, and the compatibility issues to devices, a graphical authentication system called PassMatrix is used. In PassMatrix, a password consists of only one pass-square per pass-image for a sequence of n images. The number of images is user-defined. Thus the PassMatrix is vulnerable to random guess attacks based on hot-spot analyzing. This weakness can be improved by letting users upload their own images and therefore make it more difficult for attackers to collect statistics of hot-spots. In order to protect users digital property, authentication is required every time they try to access their personal account and data. However, conducting the authentication process in public might result in potential shoulder surfing attacks. Even a complicated password can be cracked easily through shoulder surfing. Using traditional textual passwords or PIN method, users need to type their passwords to authenticate themselves and thus these passwords can be revealed easily if someone peeks over shoulder or uses video recording devices such as cell phones. To overcome this problem, they proposed a shoulder- surfing resistant authentication system based on graphical passwords, named PassMatrix. Using a one-time login indicator per image, users can point out the location of their pass-square without directly clicking or touching it, which is an action vulnerable to shoulder surfing attacks.

Andrea Bianchi, Ian Oakley, and Hyoungshick Kim [12], PassBYOP is a new graphical password scheme for public terminals that replaces the static digital images typically used in graphical password systems with personalized physical tokens, here in the form of digital pictures displayed on a physical user-owned device such as a mobile phone. Users present these images to a system camera and then enter their password as a sequence of selections on live video of the token. Improving the security of graphical password systems by integrating live video of a physical token that a user carries.

Oleg Starostenkon, Claudia Cruz-Perez, Fernando Uceda-Ponga, Vicente Alarcon-Aquino [11], they proposed straightening characters and word in CAPTCHA exploiting then a three-color bar code for their segmentation. The recognition of straightened characters and whole word is provided by the proposed original SVM-based learning classifier. The main goal of this research is to reduce vulnerability of CAPTCHA from spam and frauds as well as to provide an approach for recognizing either handwritten or degraded and damaged texts in ancient manuscripts by OCR systems. Thus, the proposed segmentation process is based on three-color bar character encoding provides satisfactory separation of letters in reCAPTCHA.

Yongdong Wu, Zhigang Zhao, Feng Bao, and Robert H. Deng [10], Denial-of-service (DoS) and distributed DoS (DDoS) are the major threats to cyber-security, and client puzzle, which demands a client to perform computationally expensive operations before being granted services from a server, is a well-known countermeasure to them. In order to prevent DoS/DDoS attackers from inflating their puzzle-solving capabilities a new client puzzle referred to as software puzzle is developed. Software puzzle scheme is proposed for defeating GPU-Inflated DoS attack. It adopts software protection technologies to ensure challenge data confidentiality and code security for an appropriate time period. It has different security requirement from the conventional cipher which demands long-term confidentiality, and code protection which focuses on long-term robustness against reverse-engineering. The software puzzle may be built upon a data puzzle, it can be integrated with any existing server-side data puzzle scheme, and easily deployed as the present client puzzle. The server has to spend time in constructing the puzzle. The present puzzle is generated at the server side. An open problem is how to construct the client-side software puzzle so as to save the server time for better defense performance.

IV. PROPOSED WORK

The proposed work focus on providing authentication and improving security. At present most of the online application use Captcha as security to prevent the spammer. The Image based Captcha allow the user to select particular images inorder to prove that the user is a human not a spammer. The algorithm used is Shuffling Chaotic Encryption Algorithm. The algorithm mainly focuses on providing confusion for the

user [9]. The images will be shuffled everytime during login and enrollment. It includes Image based Captcha as authentication that mainly focus on using homomorphic images. The geometric transformation is to be applied on selected images. The geometric transformation applied is affine transformation [8]. The user has to select the image during enrollment. The selected image will be stored into the database. During login, the user has to select the image that is selected during enrollment. If the selected image matches with the stored image, the user will be allowed to access their folder. Then the OTP is sent to registered ID for verification.

DRAWBACKS OF DIFFERENT TYPES OF CAPTCHA

A) Text Captcha- In text images, user has some problem to identify the correct text or characters. Multiple fonts, Font size. Blurred Letters iv. Wave Motion [7]. It can be easily identified by OCR techniques.
B) Audio Captcha- It is available in English therefore end user must have a comprehensive English vocabulary. Character that have similar sound.
C) Video Captcha- Due to large size of file, users face problem to download video and find correct captcha.
D) Puzzle Captcha- The task is not easy for users because puzzle based captcha take more time to solve the puzzle and identify actual arrangement of puzzles.

V. APPLICATIONS OF CAPTCHA

The various applications of Captcha are:

A) Registering the web forms
B) Online polling sites
C) E-banking
D) E-Ticketing
E) E-mail spam

VI. CONCLUSION

In this paper, the image based captcha improves the security. The Occluded image makes the difficult task for the user. Since image will appear as homomorphic images. To bring complexity for the user Shuffling chaotic encryption algorithm is being used. It overcomes the problem of several attacks. The image recognition is the task performed by the user. The recognized image is known only to the authenticated user. It provides the more challenging issue for spammer. Because of using Occluded image, it brings complexity for the spammer. Therefore, the spammer can be easily identified by the user.

REFERENCES

[1] Monica Chew, J. D. Tygar, and UC Berkeley," Image Recognition Captcha", Information Security Conference pp. 268-279, 2004.

[2] Yong Rui, and Zicheng Liu, "ARTiFACIAL: Automated Reverse Turing Test using FACIAL features", Multimedia Systems, 2004.

[3] Jeremy Elson, John R.Douceur, Jon Howell, and Jared Saul," Assira: A CAPTCHA that exploits Interest-Aligned Manual Images Categorization", Information Security Conference pp.958-67, 2007.

[4] Bin B. Zhu, Jeff Yan, Guanbo Bao, Maowei Yang, and Ning Xu, "Captcha as Graphical Passwords- A New Security Primitive Based on Hard AI Problems", IEEE Transactions on Information Forensics and Security, Vol. 9, pp. 891-904, 2014.

[5] Gaurav Goswami, Brian M. Powell, Mayank Vatsa, Richa Singh, and Afzel Noore, "FaceDCaptcha: Face detection based color image CAPTCHA", Future Generation Computer Systems 31, pp.59-68, 2014.

[6] Narges Roshanbin, and James Miller," ADAMAS: Interweaving Unicode and color to enhance CAPTCHA security", Future Generation Computer Systems 55, pp.289-310, 2014.

[7] Ved Prakash Singh, Preet Pal, "Survey of Different Types of CAPTCHA", International Journal of Computer Science and Information Technologies, Vol. 5 , pp.2242-2245, 2014.

[8] Harshit Somani, Meenu Chawla , Namita Tiwari , Madhu Shandilya, "Image Encryption using Block Shuffling and Affine Transform: A Review", International Journal of Computer Applications, Vol.95, pp.15-18, 2014

[9] Shaheen Ayyub, Praveen Kaushik, "Securing Images in Cloud using Hyper Chaos with User Authentication", International Journal of Computer Applications, Vol. No. 17, 121 – 17, July 2015.

[10] Yongdong Wu, Zhigang Zhao, Feng Bao, and Robert H. Deng, "Software Puzzle: A Countermeasure to Resource-Inflated Denial-of-Service Attacks", IEEE Transactions On Information Forensics And Security, Vol. 10, pp. 168-177, 2015

[11] Oleg Starostenkon, Claudia Cruz-Perez, Fernando Uceda-Ponga, Vicente Alarcon-Aquino," Breaking text-based CAPTCHAs with variable word and character orientation", Pattern Recognition, Vol. 48, pp. 1101-1112, 2015

[12] Andrea Bianchi, Ian Oakley, and Hyoungshick Kim, "PassBYOP: Bring Your Own Picture for Securing Graphical Passwords", IEEE Transactions on Human-Machine Systems, Vol. 46, pp. 380-389, 2016.

[13] Guido Schryen, Gerit Wagner, and Alexander Schlegel, "Development of two novel face-recognition CAPTCHAs: A security and usability study", Computers and Security pp.95-116, 2016.

[14] Hung-Min Sun, Shiuan-Tung Chen, Jyh-Haw Yeh and Chia-Yun Cheng, "A Shoulder Surfing Resistant Graphical Authentication System", Procedia Computer Science pp. 490– 498, 2016.

[15] Mohammed E. Hoque, David J. Russomanno, and Mohammed Yeasin, "2D Captchas from 3D Models", Secure Digital Information, pp. 66-72, 2016.