



OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

SURVEY PAPER ON SECURE TEXT-BASED CAPTCHA

Vaibhav Jawane¹, Prof. Vijay Rathi²

Student, Vaibhav Jawane, Dept. Computer Engineering, P.G.M.C.O.E Pune, India¹

Prof. Vijay Rathi, Dept. Computer Engineering, P.G.M.C.O.E Pune, India²

jasminjamadar@gmail.com¹

Abstract: Captcha is remains for Completely Automated Public Turing test to distinguish Computer and Human One from the other. As the expansion of robotized bots frameworks or programming that abuse and degenerate the general population web benefits, the client must required experiencing and settling a Turing test issue, before they are utilize web administrations. This Turing test is called Captcha. In this paper we have examine an enhanced content based captcha which is more secure, and more hearty when contrasted with another Captchas.

Keywords – Captcha ,Optical character recognition (OCR).

I INTRODUCTION

Now a day's everybody need web application. As the internet growth, its security is also becoming critical issue. There have been lots of abuse system and bot programs the corrupt and affect the web services. That is way websites uses Captcha to differentiate human and computer program apart 1 CAPTCHAs come in various forms. Mostly text-based Captcha is used that simply an image format that contains alphanumeric characters that placed in front and distortion or degradation apply background of these images to make optical character recognition (OCR) cannot recognize these images 2 . A good quality captcha is robust and secure. A variety of CAPTCHA have been proposed

II RELATED WORK

We concentrate on Text-based Captcha in light of the fact that it's for the most part utilized; in light of the fact that effectively actualize and minimal effort .Optical Character acknowledgment (OCR) is utilized to break content based captcha. As a rule OCR programs perceive characters in a picture through three steps: (1) pre-handling of the picture to make the picture reasonable for additionally preparing, (2) dividing the picture into areas in which every district contains just a single character, and (3) recognizing the character in every district. To bring down the achievement rate of character acknowledgment by the OCR programs, CAPTCHA frameworks more often than not misshape the pictures in certain approaches to muddle the

means OCR programs commonly utilize. As the OCR assaults are solid there is have to solid the content based captcha so clamor included, utilize hued foundations, utilize of different shapes, for example, thin lines, thick lines, circles and rectangles etc.

2.1 Literature Survey

In this paper, a new type CAPTCHA system will be proposed. The proposed scheme, named Clickspell, combined the features of text-based and image-based CAPTCHAs[1]. Click spell, tries to increase each security and usefulness of CAPTCHA, Clickspell asks users to spell a randomly chosen word by clicking letters for passing the test all the letters are distorted. This paper discusses usability issues that should be considered and addressed in the design of CAPTCHAs. Some of these issues are intuitive, but some others have subtle implications for robustness (or security). A simple but novel framework for examining CAPTCHA usability is also proposed.[1]

As the increase of automated bots systems or software that misuse and corrupt the public web services, the user must required going through and solving a Turing test problem, before they are use web services[4]. This Turing test is called Captcha. In this paper we have discuss an improved text-based captcha which is more secure, and more robust as compared to another Captchas.[4]

The massive and automated access to Web resources through robots has made it essential for Web service providers to make some conclusion about whether a "user" is human or robot.[6] A Human Interaction Proof (HIP) like

Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) offers a way to make such a distinction. CAPTCHA is essentially a modern implementation of Turing test, which carries out its job through a particular text based, image based or audio based challenge response system. In this paper we present a new image based CAPTCHA technique. This paper we have discuss an improved text-based captcha which is more secure, and more robust as compared to another Captchas[9].

In this paper propose a novel framework that guides the application of our well-tested security engineering methodology for evaluating CAPTCHA robustness, and we propose a new general principle for CAPTCHA [10]

III METHODOLOGY OF PROPOSED WORK

We have proposed an enhanced content based captcha which is unbreakable, more secure, and more powerful as contrasted with existing Captchas. We have proposed another plan of a Captcha which is a mix of haphazardly produced Mathematical arithmetical condition and an alphanumeric word. Distinctive situations containing change of Textual Style, change of Alignment, change of position, arbitrarily created words have been used to check the helplessness of proposed captcha against known discovery strategies. The outcomes acquired are extremely surprising as contrasted with the current style of Text based Captchas.

CAPTCHA's code is a series of characters (uppercase and lowercase) and numbers. Multiple randomizing functions are used to generate a random code (stream of characters and numbers) in each challenge in order to make it not susceptible to a dictionary attack. The length of the code is varied (minimum length is 6 characters-numbers). Multiple font types are handled to prevent intrusion using image processing techniques when a consistent font is used. The text image is blurred using a specific technique in order to make CAPTCHA difficult for malicious software.

3.1 Steps for the Proposed Model

1. Select the random digits and alphabets.
2. Make a string that has mathematical algebraic equation using d and an alphanumeric word.
3. Generate an image and embed the above generated image on it.
4. Repeat the steps to 1, 2, 3 to make database of images.
5. One random image from data base is given to user. User fill the text that is shown is given image.
6. Then text fill by user is check with database if matches then user pass the test.
7. Along with this system captcha security also check using OCR.

3.2 Proposed captcha

We focus on Text-based Captcha because it's mostly used; because easily implement and low cost .Optical Character recognition (OCR) is used to break text-based captcha4 . In general OCR programs recognize characters in an image via three steps : (1) pre-processing of the image to make the image suitable for further processing, (2) segmenting the image into regions in which each region contains only one character, and (3) identifying the character in each region. To lower the success rate of character recognition by the OCR programs, CAPTCHA systems usually distort the images in certain ways to complicate the steps OCR programs typically employ A captcha is a combination of randomly generated mathematical algebraic equation and an alphanumeric word shows in figure 1. Generation of image containing different scenarios:



Figure 1: Shows the proposed captcha.

- 1) Change of font style: In earlier Captch as single font is used for every character. But in our proposed captcha there is multiple fonts is for every character.
- 2) Change of Alignment: In our proposed captcha alignment of character is vary. They are not in a same line.
- 3) Change of position: Every time when new captcha is generated the position of mathematical equation and alphanumeric word is varying means may be alphanumeric equation is at starting and mathematical algebraic equation at end.
- 4) Randomly generated words: We are not using dictionary words because they are easily breakable.

3.3 Optical character recognition(captcha assault)

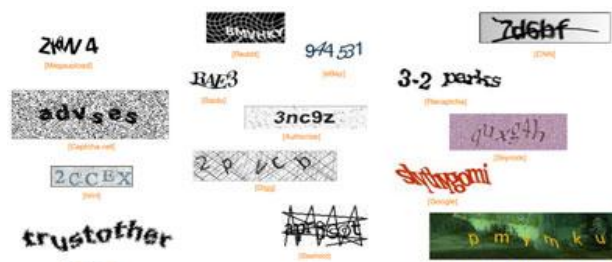


Figure 2: Sample of Captchas.

Optical character acknowledgment is a procedure that is utilized for breaking the captcha. There are for the most part three stages of OCR : preprocessing, division and acknowledgment . The pre-handling stage is first phase of OCR that endeavors to clean foundation commotion, stray lines, and other pointless subtle elements. Picture is parallels with the goal that lone vital data is passed for additionally handling. Division is most troublesome errand in OCR. Picture is soften up locales. These districts might be letters, words or whatever other locale that will be handled further. The last stage – acknowledgment includes recognizing the area and arranging it as one from the space of conceivable answers.

IV RESULTS

We take the database of 100 images and apply optical character recognition test on them we can see only 5times proposed system is break. The success rate of proposed system is 95%. We compare our proposed captcha with earlier captchas shows in figure 2 . Results of different Captchas show in table 1.

Table no 1-Results of different Captchas

Sr. no	Name of captcha	Percentage of breaking
1	Gimpy	81%
2	EZ-gimpy	92%
3	Securimage	100%
4	Cryptography	100%
5	Mega upload	63%
6	New proposed system	4%

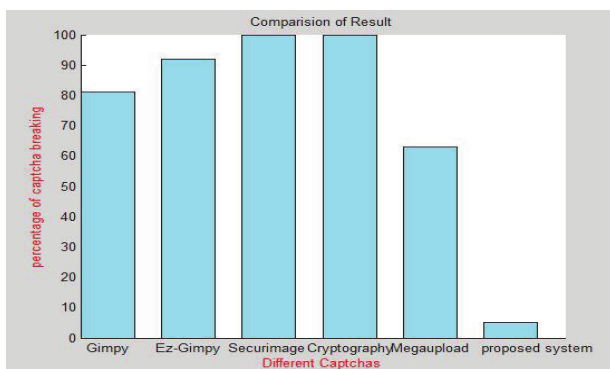


Figure 3: Comparison of different captcha

V CONCLUSION

Proposed model is text-based captcha that have various scenarios of text. To check the vulnerability of proposed captcha OCR is used. This technique is easily

understandable and secure. The proposed system is in progress phase. Combination text and audio system will be implementing as a future work. The proposed system provides great security as compare to earlier Captchas.

ACKNOWLEDGEMENTS

I am short of words to express our thankfulness to all the distinguished people who during my work gave me their unflinching support and valuable guidance, which helped me to make my endeavour a success. It is my pleasure to acknowledge **Mr.Vijay Rathi** with deep sense of gratitude for their valuable time and help; they devoted on me and enlightened me with lot of knowledge for successful start and proper continuation of this project.

REFERENCES

- 1.Prof. Yogdhar Pandey and Darshika Lothe, “Evaluating the Usability and Security of a Spelling Based Captcha System,” International Journal of Computer Science and Information Technologies, Vol. 5 (3), 2014, 4728-4731.
2. Clark Pope and Khushpreet Kaur, “Is It Human or Computer? Defending E-Commerce with Captcha,” IEEE Computer Society, March – April, IT PRO 2005.
3. J. Yan, A.S.E. Ahmad, “Usability of CAPTCHAs or usability issues in CAPTCHA design,” Presented at the 4th Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania, 2008.
4. Chih-Hsiang Hsu and Ying-Lien Lee, “Usability study of text-based CAPTCHAs,” 2011Elsevier, vol.32 (2011) 81–86.
5. J.Ogijenko and A.Kolupaev, “Captchas: Humans vs Bots.” IEEE Computer Security, vol.6, pp.68-70, Feb. 2008.
6. M. Tariq Banday and Nisar A. Shah, “Drag and Drop Image CAPTCHA.” Sprouts 4 th J&K Science congress, 2008 .
7. Ahmad El Ahmad, Jeff Yan and Wai-Yin Ng , “CAPTCHA Design Color, , and Security,” IEEE Computer Society, 1089-7801, March – April 2012 IEEE.
8. N. A. Shah and M. Tariq Banday, “A Study of CAPTCHAs for Securing Web Services.” IJSDIA International Journal of Secure Digital Information Age, Vol. 1. No. 2, December 2009.
9. M. Chew and J. D. Tygar, “Image Recognition CAPTCHAs”, In Proceedings of the 7th International Information Security Conference (ISC 2004). 2004, Springer.
10. R.Bazzi, R.Rusn and L.J.Karam, “A captcha based on the human visual system masking characterizes.” ACME, 2006.
11. Ahmad El Ahmad and Jeff Yan, “Colour, usability and security: a case study.” computing science Newcastle University, May 2010.