

# Survey of Human Interactive Proof and CAPTCHA Techniques and Attacks

<sup>1</sup>Ashima Narula, <sup>2</sup>Swati Bhasin

<sup>1</sup>Student

<sup>1</sup>Electronics and Communication Engineering,

<sup>1</sup>Geeta Institute of Management And Technology, Kurukshetra, India

**Abstract** - Credit division and optical character (OCR) technologies with the development, skills gap and deformed humans and bots in acts between related information becomes increasingly smaller. This trend should eventually ineffective potentially drawing API Text CAPTCHAs. To arrange these obfuscations open test, how to interpret them graphically display the image that users have plenty of them superficial effort to prevent a resolution are not yet within reach to solve the employment of computer vision algorithms. To captcha challenge to make a hard question for us to understand who pass phrase set the challenge. And several plans have been implemented. A special design for user friendliness based on results of a survey conducted on human subjects was determined. The survey result shows that our design is preferred by users and automated using known methods helps to prevent attacks from spambots. The letters Captcha technology and human interactive proof and offers a survey of attacks.

**Keywords** - CAPCHA. Security, CAPTCHA attacks; OCR, Computer vision.

## I. INTRODUCTION

Slowly increase the interactivity of complex web requests for the appointment of a term that has come to be. Within this highly vibrant and evolving context, conservation issues and concerns have recently upgraded to the center of attention. Indeed, after a security breakage can damage the credibility and responsibility of a union is valid, users manage to defeat? Trust. In this context, an important security concern is primarily that malicious users or agents multimedia system resource as a party to the general deterioration of the quality of the result of misuse of the ability for users to automate effort. Fully Automated Turing test area to inform Computers and Humans apart (CAPTCHA) a challenge-response test to protect against such threats is like. To protect these tests that it is a human attempt to interact with a web system. CAPTCHA tests often distorted images embedded in the text, knowing the kind of man-made intelligence failures are set to open. These open far too early failures that cannot be solved by a computer, but the rash can be solved by a human. Two important qualities of a good CAPTCHA dimensions are its safety and usability. A CAPTCHA security level determines its strength in challenging adversarial attacks. In the same period, a request for a web application to protect the CAPTCHA casual user interacting with a pair of goals to reduce the cognitive power, it should be more user-friendly and transparent. CAPTCHA check on significant attention recently consented. A lot of work, mixed CAPTCHA system is focused on the protection generally on how to break and how to raise them. And implementation of work widely used to assess the association and to enhance their security or usability is focused on methods of consultation. Recently, non-text CAPTCHA mechanism, like audio and video is an increased focus on. In each case, it is accepted that the CAPTCHA system design far more an art than science. The purpose of this paper beliefs, customs, and user preferences CAPTCHA connected are connected to the user to check the sights. Human Interactive Proof (HIPS) tests for a user that can be delivered through multimedia assure that as a human being challenged in an automated system, is negotiating with the software to help are. Hips, being used more and more to prevent the web from places, for example, many Web-based email accounts automatically conception. Usually such reports are automatically generated spam aggressiveness and are used for supplementary non-desirable activities. A picture that is a public example of a hip text that can be a real word or phrase, or message, points, and may include supplementary characters nonsensical combinations. Hip test, is shown in acts like a user to solve. Tests (e.g., audio and / or video challenges) Supplementary type additionally, the hips can be as industrial intend to find out, for example, whether a particular appeal approved by a web locale Start by being a human being with. While a small unit (e.g., large corporations) have the resources to investigate and hip testing an automated system that are difficult to solve, can develop, countless supplementary entities (e.g., small web based businesses), hip able to use the knowledge resources, the challenges of developing the desire but could not. A CAPTCHA (Completely Automated Turing test area to inform Computers and Humans Apart) is a plan that generates and grades tests that are human solvable, but current computer programs beyond the skills [1]. Nowadays this knowledge (such as junk email is spreading and now thousands of free email report them as grasping) solution for unwanted or malicious Internet bots plans average about a safety net and countless business web Google, Yahoo added places discovery requests, and Microsoft's MSN.



FIGURE1:CAPTCHA

The word "captcha" Vaughan et al in 2000, was an early. [1] Is a test that can distinguish humans from computers to describe? Public indentation down, be tempted

1. Facilely resolved by humans,
2. Facilely generated and evaluated,
3. But, not facilely resolved by computer

Over the past decade, captchas industrialized production has a number of different ways, each one painted quality to meet the fluctuating degrees. CAPTCHAs most commonly observable tests showed that users need sound and distortion of an image by combining a little vague in identifying alphanumeric acts are. Figure 1 displays are examples of such observable CAPTCHAs. So far, there are three main types of CAPTCHAs pursuing:

1. Text-based schemes – they generally rely unsophisticated distortion of text pictures interpretation them unrecognizable to the state of the fine art of pattern recognition plans but recognizable to human eyes.
2. Sound-based schemes (or audio schemes): - they typically need users to resolve a speech recognition task.
3. Image-based schemes - they normally need users to Performa picture credit task.

In this paper, our discussion will focus mainly on the text-basedCAPTCHAs, following reasons:

First, the text-based CAPTCHAs have been most extensively used plans. Such as Google, Yahoo and Microsoft as the main web locations have their own text-based CAPTCHAs used for years.

Second, the text-based CAPTCHAs countless benefits schemes [4] than the type of supplement, for example, is intuitive to users world-wide (the user is given the task of being the only character recognition) insufficient to localize themes, and good ability to offer strong protection (such as an animal force attack space, can be immense if properly designed've got).

Third, it is a matter that should be addressed in these plans recognizes such claims are accepted and well captchas to improve the usability of the area can be a huge and positive encounter. Arranging these obfuscating openly test (vague, obscure, or to make unintelligible) CAPTCHAs superficial plenty of them attempt a resolution that did not stop the (discouraging) barely within reach to make computer vision algorithms to solve the employment, yet. The mismatch between the current knowledge, though slimmer and narrower. It is strategically sound and adds supplementary distortions, or to arrange an extra tightly acts is possible to enhance the security of continuous text CAPTCHA. These measures, however, difficult man to understand the functions of addition, a high error rate and high web load should emerge. Credit division and optical character (OCR) technologies with the development, skills gap and deformed humans and bots in acts between related information becomes increasingly smaller. This trend should finally painting text CAPTCHAs potentially ineffective API.

## II. TYPES OF CAPTCHA

### A: Image established CAPTCHA

Picture Credit Line established the task of presenting a CAPTCHA users demand. CAPTCHA is believed to make more difficult to understand by bots is to use pictures. Credit to all the pixels in color as well as meaningful to the vast collection of images is extremely difficult in the CAPTCHA. This need for user-submitted images easy to recognize objects. Picture identification CAPTCHA, human emotion set CAPTCHA, pictures of landscapes set up CAPTCHA come down. Picatcha the picture that they are asked to identify a selection provides the user with a primary election. A statement or a set of human emotion Captcha graphic user is displayed. Such as an answer to the user in a thread describes their feeling. CAPTCHA is installed behind the scenario considered only knowing of objects instead of human beings is to use analytical and comprehension skills. Photo CAPTCHA established, we consider the features of interactive captcha attention credit CAPTCHA.

### B: Interactive CAPTCHA

CAPTCHA user through a sequence of interactive dialogue to solve a CAPTCHA test a user's needs. In this, a little background clutter up a general picture of vibrantly CAPTCHA is generated and displayed. The user clicks on the image CAPTCHA. The CAPTCHA image is clicked; with obscure acts materialize countless buttons CAPTCHA image below. Users in the Next button CAPTCHA image are to quickly click on the corresponding character. Click on each one, has provided a new set of buttons. The input sequence ever a click for every single character has been CAPTCHA image. Correct answers on the server side session data is stored in the index and the user clicks concerning the indices. After finishing the input sequence, is contrasted with the correct index sequence. If there is a match, CAPTCHA has been decided by the user. Interactive CAPTCHA security measures for the

duration it takes away a user on a per-character basis for submission to answer. Therefore, interactive CAPTCHA worth average per-character timeout can be set lower than that timeout. The discovery of human aggression, because every time an input and to dispatch a human solver seizes the CAPTCHA is modest in the comparative period provides a much greater resolution. Interactive quick to decode CAPTCHA images as the first multi-step challenge / response sequence to start the period as users seeking to seize allows. The limitation of interactive CAPTCHA is not suitable for blind people.

### III. PROBLEMS WITH CAPTCHAS

CAPTCHAs have several limitations, including:

A: Usability is always an important issue in designing a CAPTCHA. CAPTCHA utility and robustness are two basic issues, and they are often at loggerheads with each other. CAPTCHA only known work is focused on addressing the usability aspect of the design that should be recognized CAPTCHA human-friendly”.

B :We Can enhance the security of an existing text CAPTCHA by adding noise and distortion and arranging characters more tightly.

C: But There is a limit to the distortion and noise that humans can tolerate in a challenge of a text CAPTCHA.

D: These measures, however, would also make the characters harder for humans to recognize, resulting in a higher error.

### IV Why are the image-based CAPTCHAs NOT so extensive as the text-based ones?

I crack them conflicting scrutiny on potential stroke, you have a little consideration / calculations can get here and here. As a web developer I think the desire to express my point. So why? They are too big. A CAPTCHA on a web page should not seize a dominant position. Just fill in the form at the bottom of an auxiliary agent that serves to clear bots, becoming a small data, etc. traffic. An inadequate images, 5-10 KB each one with a size of the subject, a lot of a single page, in my opinion should weigh. Employing a low-bandwidth web visitors will be unpleasantly surprised, employment dial-up connection to create a website for visitors to comment, not conception anomaly. Cats (or dolphin) with a CAPTCHA on a vacation locale would be appropriate, but will be irrelevant - for example - in a place of a health institution. In this case, the health topics is possible to gather a number of pictures, but - Asirra on the analogy - a large number of photographs of the homeless with doctors to find a place to be problematic center of the picture painstaking The process of creation smile.

### V Attacks on Text Based CAPTCHAS:

A history of how CAPTCHA has been adopted in excess of the years comes out to be informative. Most of sites adopted CAPTCHA because their resources to send spam or anonymous, by the motive of conducting illegal activities were easy to abuse. Consequently, captchas are used widely than before, which is common part of the current website login system. However, CAPTCHA implementation is arduous and insecure idea without design. With some methods, your website can be easily cracked CAPTCHA scheme [3].

- Text-based CAPTCHAs are definitely less "secure" than their image/audio counterparts.
- As it is more accessible to impaired ones , logic questions require acumen than image CAPTCHAs.
- Neural Networks mainly used to interpret text based CAPTCHA.
- Heuristic checks: Heuristics are exposition in a process that seem to indicate a given result. It also possible to detect the presence of a robotic user based on the volume of data the user requests, series of common pages visited, IP addresses, data entry methods, or other signature data that can be collected.
- Online CAPTCHA Services: online CAPTCHA breaking services had made it possible to rupture CAPTCHA without any effort by the user.

### VI ACCESSING SECURITY ON GIVEN CAPTCHA API

Captcha and supplementary CAPTCHA API ability providers validate millions of CAPTCHAs every single date and protect thousands of websites opposing the bots. A safeguard CAPTCHA creation and validation ecosystem forms the basis of the public belief ideal amid the CAPTCHA provider and the consumer. A collection of damage can transpire if each constituent of this ecosystem is compromised. CAPTCHA API providers usually proposal both CAPTCHA creation and validation services. To consume these services, the subscribing websites whichever use the continuing libraries and plug ins; or comprise their own. A normal user contact alongside a web request that relies on a CAPTCHA provider is summarized below:

1. A user requests a page that requires CAPTCHA confirmation
2. The back up page consist an embedded <img> (or <script>) tag to find out the CAPTCHA image from the user who provide CAPTCHA .

3. Upon parsing the embedded tags, the browser find out a CAPTCHA from the CAPTCHA provider and displays it to the user.
4. The user fills in the form fields, enters the CAPTCHA solution and submits the page to the web application.
5. The web application then comply the CAPTCHA solution to the CAPTCHA provider for certification
6. The CAPTCHA provider acknowledge to the web application with success or failure message.
7. Based on CAPTCHA provider's response, the web application accept or reject the request.

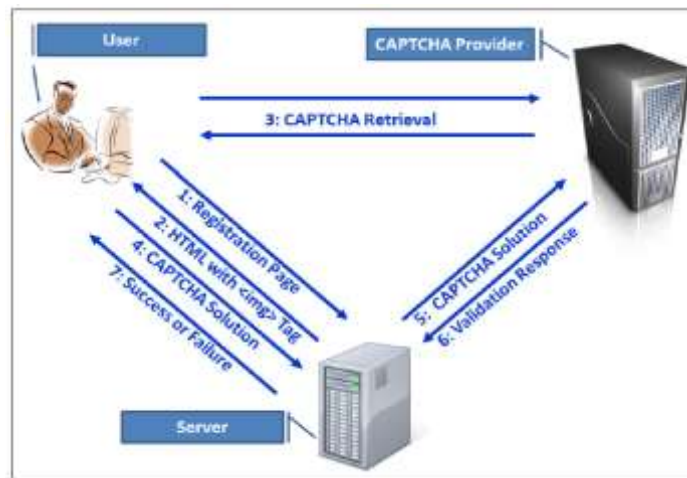


Figure 2: A typical validation flow with CAPTCHA providers

Steps 5 and 6 play a important role in the CAPTCHA validation scheme and must be securely implemented to prevent attacks against CAPTCHA validation process.

## VII OPEN ISSUES ON CAPTCHA DESIGN

In order to eliminate the possibility of a computer attack, there is still a lot we could do. CAPTCHA can be further improved: We can enhance human participation and make full use of our human brain, such as simple mathematic test, sentence and image understanding, which could hardly have learned by computer vision. We could use human other senses except visual decryption system and sound filtering system for CAPTCHA application. Current CAPTCHA have not realized how complicated human brain is. Other human senses, including touch, smell, quick learning or logical analysis ability, did not be exploited by CAPTCHA designer. We could create new form [11] for CAPTCHA application. With the development of the media technology, mobile phones and smart televisions are competing for users with computer. Creating CAPTCHA for other forms of media becomes necessary than before. Then other media would be protected by authentication.

## VIII CONCLUSION AND FUTURE SCOPE

CAPTCHAs actions that degrade the quality of a given system's ability to deliver automated multimedia are used in attempts to prevent. CAPTCHAs addition mixed sites are used to reduce the automatic posting. Security. It systematic distortion sound and add useful, and the system acts more tightly than is possible to increase the security of a continuous text CAPTCHAs CAPTCHA countless supplement is requested. These measures, however, difficult man to understand the functions of addition, a higher error rate should emerge. There is distortion and sound that humans can tolerate a text CAPTCHA for a test is a check. Usability is always an important topic to arrange a CAPTCHA. Credit division and optical character (OCR) technologies with the development, skills gap and deformed humans and bots in acts between related information becomes increasingly smaller. This trend should eventually ineffective potential painting text CAPTCHAs. This suggests that the CAPTCHA text level work in an overtly aggressive. Aggression mixed OCR text CAPTCHAs to be applied to a finished relationship materializes. Aggression to protect the character of the crowd jointly accepted division immune system is at the top of the craft. CAPTCHAs are still a new investigation period, open failures mislabeling occurs. They argued all the setbacks, mislabeling causes the human errors. The authors, this may be able to resolve the employment cooperative filtering recognized human users rate pictures according to how well their label called while. By design, Text CAPTCHAs are easy and facile to resolve by humans. Their low encounter quality makes them appealing to locale operators who are distressed of each protection that might coil away possible visitors. Though, this alike quality has made them facile to attack. In this thesis, we have debates the resolving CAPTCHAs employing Open Basis OCRs, Displaying that CAPTCHAs vulnerable to such attacks. A lot of work has been completed in Enhancing CAPTCHA usability and Protection one such example is use of CAPTCHs, though rise of present adverts and methods made it extra tough to stop automated bots and supplementary hazardous spammers opposing CAPTCHA attacks. a little method we have debated in this thesis furnish extra than 40% accomplishment rate, and as the defective CAPTCHA demands are re-evaluated by the server and nonexistence manipulating count way that CAPTCHA decryption will be prosperous in consecutive attacks. In upcoming we

should like to use our methods to more enhance the accomplishment rate of the CAPTCHAs. One more span of work will be to enhance picture heuristics. For enhancing intricacy of the OCR for credit of the words, a background picture will be added to the image. This background picture could encompass of a little random lines or a little noise. Removing these lines is a hard task for an OCR system. Because removing these lines could obliterate a little spots of the messages and change a character to another. Also cursive character credit, the segmentation of the acts is harder than knowing the characters. So, adding lines make the segmentation of acts a tough for an OCR program

## REFERENCES

- [1] Saxena, Ashutosh, Nitin Singh Chauhan, K. R. Sravan, Aparajith Srinivasan Vangal, and David Palacios Rodriguez. "A new scheme for mobile based CAPTCHA service on Cloud." In *Cloud Computing in Emerging Markets (CCEM)*, 2012 IEEE International Conference on, pp. 1-6. IEEE, 2012.
- [2] D'Souza, Darryl, Phani C. Polina, and Roman V. Yampolskiy. "Avatar captcha: Telling computers and humans apart via face classification." In *Electro/Information Technology (EIT)*, 2012 IEEE International Conference on, pp. 1-6. IEEE, 2012.
- [3] Kiran Jain Azad, "CAPTCHA: Attacks and weaknesses against OCR Technology." *Global Journal of Computer Science and Technology* 13, no. 3 (2013).
- [4] Achint Thomas, KunalPunera, Lyndon Kennedy, Belle Tseng, and Yi Chang. "Framework for evaluation of text captchas." In *Proceedings of the 22nd international conference on World Wide Web companion*, pp. 159-160. International World Wide Web Conferences Steering Committee, 2013.
- [5] Ye, Quan-Bin, Te-En Wei, Albert B. Jeng, Hahn-Ming Lee, and Kuo-Ping Wu. "DDIM-CAPTCHA: A Novel Drag-n-Drop Interactive Masking CAPTCHA against the Third Party Human Attacks." In *Technologies and Applications of Artificial Intelligence (TAAI)*, 2013 Conference on, pp. 158-163. IEEE, 2013.
- [6] Men, Tao, Deming Wang, Yan Sun, and Mingrong Wang. "A novel dynamic CAPTCHA based on inverted colors." In *Instrumentation and Measurement, Sensor Network and Automation (IMSNA)*, 2013 2nd International Symposium on, pp. 796-799. IEEE, 2013.
- [7] Yunhang Shen, Rongrong Ji, Donglin Cao, and Min Wang. "Hacking Chinese Touclick CAPTCHA by Multi-Scale Corner Structure Model with Fast Pattern Matching." In *Proceedings of the ACM International Conference on Multimedia*, pp. 853-856. ACM, 2014.
- [8] Max Jaderberg, Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. "Synthetic data and artificial neural networks for natural scene text recognition." arXiv preprint arXiv:1406.2227 (2014).
- [9] Min Wang, Tianhui Zhang, Wenrong Jiang, and Hao Song. "The recognition of CAPTCHA." *Journal of Computer and Communications* 2, no. 02 (2014): 14.
- [10] Haichang Gao, Wei Wang, Ye Fan, Jiao Qi, and Xiyang Liu. "The Robustness of "Connecting Characters Together" CAPTCHAs." *Journal of Information Science and Engineering* 30, no. 2 (2014): 347-369.
- [11] AriyanZarei, "Improve CAPTCHA's Security Using Gaussian Blur Filter." arXiv preprint arXiv:1410.4441 (2014).
- [12] Zhu, B., Jeff Yan, GuanboBao, M. Mao, and Ning Xu. "Captcha as Graphical Passwords—A New Security Primitive Based on Hard AI Problems." (2014): 1-1.
- [13] Carlos Javier Hernández-Castro, David F. Barrero, and María D. R-Moreno. "A Machine Learning Attack against the Civil Rights CAPTCHA." In *Intelligent Distributed Computing VIII*, pp. 239-248. Springer International Publishing, 2015.