International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

CAPTCHA SECURITY ENGINE

G.DHANALAKSHMI [1]S.DEVADARSHINI [2],R.SWARNAA SRI [3],R.SRINIDHI [4]

[1]*Assistant Professor, Department of Information Technology, Panimalar Institute of Technology, Chennai, India*

[2,3,4]*Student, Department of Information Technology, Panimalar Institute of Technology, Chennai, India.*

Abstract : *The CAPTCHA is employed to supply the safety against the malicious software by generating a test which only a person can complete. The CAPTCHA stands for completely automated public Turing test to inform computer and human apart. Currently, we are using CAPTCHA are image and text-based data. The CAPTCHA security engine will provide a far better thanks to generating the data for CAPTCHA and can increase the difficulty in bypassing the system by use of improved algorithm.*

***Key Words*: Convolutional Neural Network, Deep Learning, Forest fire, Loss Function, Machine Learning, Optimization Algorithm, YOLOv3.**

1. INTRODUCTION

CAPTCHA can be deployed to protect systems vulnerable to e-mail spam, such as the webmail services of Gmail, Hotmail, and Yahoo!. CAPTCHA have also found active use in stopping automated posting to blogs or forums, whether as a result of commercial promotion, or harassment and vandalism. CAPTCHA also serve an important function in rate limiting, as automated usage of a service might be desirable until such usage is done in excess, and to the detriment of human users. An example of a system in which vulnerabilities exist, which could easily be prevented using CAPTCHA. CAPTCHAs are used to prevent bots from using various types of computing services. Applications include preventing bots from taking part in online polls, registering for free email accounts (which may then be used to send spam), and, more recently, preventing bot generated spam by requiring that the (unrecognized) sender pass a CAPTCHA test before the email message is delivered. They have also been used to prevent people from using bots to assist with massive downloading of content from multimedia websites. CAPTCHAs are used in online message boards and blog comments to prevent bots from posting spam links as a comment or message. CAPTCHA is a question/answer based CAPTCHA system. It has been designed to protect any Web form from spam bots attempting to submit information automatically by asking questions the visitor needs to answer. It can also detect brute force attempts and will block further abuse by adding the visitor's IP address to the .access file. CAPTCHA prevents accidentally blocking good bots by verifying, who is record of the visitor's IP. It is also possible to use CAPTCHA as an automatically blocking bot trap. Bot trap implementation guide can be found in the guides.

EXISTING SYSTEM

In the existing CAPTCHA security system we use mainly text based CAPTCHA which we are using from the begging of time, so the today there is some software which can bypass this test. The hacker hacks the info from the system then they create software consistent with that and therefore the CAPTCHA is bypassed. While a person's got to enter an extended sentence within the box before the access to the web site . While solving the CAPTCHA is a boring and sometimes even though it is right, it shows error and we can say that at first CAPTCHA use to protect from spam bot, but today bots are defeating the CAPTCHA while sometimes humans can't solve it. Another reason why CAPTCHA is seen as difficult to read for computers is its visual component. Because the symbols are in an image format, it's more difficult for computers to scan an image with text, especially when the text is distorted. Humans can check out a picture and detect patterns more easily.

PROPOSEDSYSTEM

In the proposed CAPTCHA security system will generate the CAPTCHA by employing a new improved algorithm which can be interesting tom was unravel at an equivalent time it'll be tougher than previous to solved by the bots while will feel easy. The humans have limitations on the speed of response then compared to any computer and hence the pc must be slower than the human then we'll make full benefit of this and use it in the proposed system. The CAPTCHA security system will contain colored graphical interface with the font is limited to two while the border line thickness and color will be fixed. The CAPTCHA security system will generate random text which can be shorter than this system but are going to be difficult to hack because it will randomly generate. CAPTCHA shows a vital starring role in the life of safety wherever it precludes the Bot hackers from abusing the web service area. This methodology proposes the CAPTCHA on cloud and provides the scheme to be safe and less defenseless. Recognizing objects in an over-distorted test is difficult or impossible for humans. In the proposed CAPTCHA, the availability of background noise which can shape, color or location can potentially baffle attackers. In order to improve the

usability and to help human users to recognize target objects in the noisy background, they will be provided with zooming and color-filtering tools. A strategy to make CAPTCHAs stronger is imposing restrictions on the number of CAPTCHAs a user can try, the maximum number of attempts to solve a test, and the duration of the validity of a test . If an IP address tries to download a large number of CAPTCHAs, it might be a robot trying to collect the CAPTCHA database. If an IP address performs several attempts on a test or spends a lot of time solving a test, it is probably a computer program trying to solve the test by random guessing or image processing attacks. Such IPs can be deprived of the service after a maximum number of attempts or at a given time.

.

II.OPERATIONS

Scanned text is subjected to analysis by two different OCRs. Any word that is deciphered differently by the two OCR programs or that is not in an English dictionary is marked as "suspicious" and converted into a CAPTCHA. The suspicious word is displayed, out of context, sometimes along with a control word already known. If the human types the control word correctly, then the response to the questionable word is accepted as probably valid. If enough users were to correctly type the control word, but incorrectly type the second word which OCR had failed to recognize, then the digital version of documents could end up containing the incorrect word. Those words that are consistently given a single identity by human judges are later recycled as control words. If the first three guesses match each other but do not match either of the OCRs, they are considered a correct answer, and the word becomes a control word. When six users reject a word before any correct spelling is chosen, the word is discarded as unreadable. The original reusable CAPTCHA method was designed to show the questionable words separately, as out-of-context correction, rather than in use, such as within a phrase of five words from the original document. In 2012, reusable CAPTCHA began using photographs , questions, puzzles.

III. HARDWARE AND SOFTWARE COMPONENTS USED

Software requirement

- Operating system : windows XP or higher
- Languages : Core Java, Servlets, JSPs, JDBC, TML
- Data Base : Oracle 10g
- Tools : Eclipse / Net Beans

Hardware requirements

- Processor : Pentium 3 with 600 MHz or above
- Ram : Minimum 1GB
- Hard disk : Minimum 30GB

IV.RANDOM GUESSING ATTACKS

In random guessing attack, also mentioned as blind guessing or no-effort attack, an attacker tries to interrupt a CAPTCHA by guessing the solution . In text-based CAPTCHAs, given the list size, c, the probability of solving an n-character CAPTCHA challenge by blind guessing is 1 /Cn. In an image-based CAPTCHA that asks a user to detect an object between n candidate objects, the likelihood of an accurate guess is 1 /n. Weaknesses of a CAPTCHA that can make it vulnerable to this attack include using a small input space, having a small number of candidate objects in a test, and imposing no limits on the amount of attempts to unravel a test.

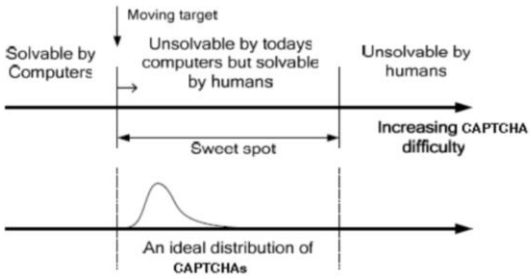
V.SOCIAL ENGINEERING TO BREAK CAPTCHAS

Using cheap third party humans is a way to break CAPTCHAs. Kang et al designed a CAPTCHA phishing attack that is a form of social engineering attack. They deployed a CAPTCHA phishing interface on a webpage (called CAPTCHA carrier) and selected some high traffic website as phishing areas to publish their phishing messages. Phishing carrier and phishing area can be two different webpages. Alternatively, phishing carrier can be integrated into the phishing area using Adobe Flash. Since people are less willing to click an unknown hyperlink, the second approach has been more successful. Troung et al designed another attack called “Instant Messenger CAPTCHA attack”. The major components of this attack are an attack script and an IM connector. The first component scrapes CAPTCHA images and uses the IM connector, to send them to the 3rd party human who solves the tests. Since instant messengers allow a real-time communication between participants, the attack cannot be detected using timeout values. The breakdown of the research work is completed in this segment and is tabularized in Table 1 In Table 1, the % attainment rate denotes that out of 100 attacks done on a CAPTCHA the assessment accompanying with it is the quantity of times the CAPTCHA is “broken” positively. It has been frequently rummage-sale because of easily manageable on the internet. The CAPTCHA execution is extremely dangerous without

suspicious scheme .There is various bouts on script constructed CAPTCHA system. Most of end result from unfortunate safety and uncaring proposal superintend. For this persistence they provide good and well safety to avoid from occurrences.

VI. ROBUSTNESS OF CAPTCHAS

A good CAPTCHA must be both easy-to-solve for humans and powerful enough to resist attacks. Designing CAPTCHAs that fulfil both usability and robustness criteria is difficult. The key point to design such a CAPTCHA is to exploit the gap in the recognition abilities between humans and computers

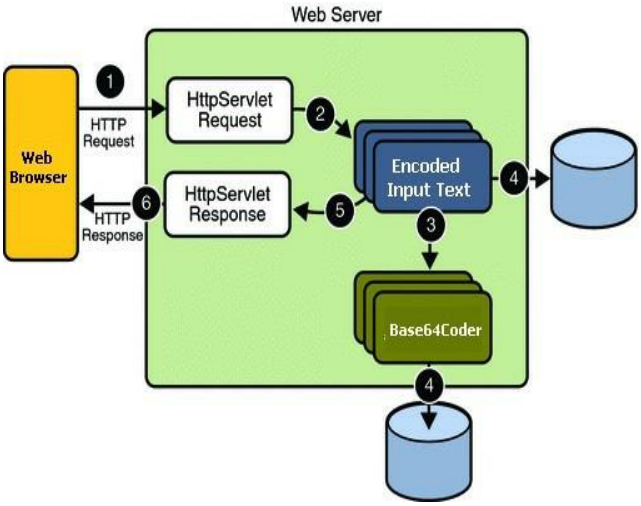


A CAPTCHA method is called strong if no automated computer program can solve its challenges with a success rate of higher than 0.01% . The security of a particular CAPTCHA test can be analysed by investigating its resistance to attacks that possibly may be used to break it. We divided this study into five subsections: - Segmentation resistance, - Recognition Resistance, - Random guessing resistance, - Security against 3rd party human solver attack; and - Other security measures. Most techniques are related to text-based CAPTCHAs because this type of CAPTCHA has been studied more than others. Many of the strategies might be mentioned in one category, but be applicable to other categories as well.

VII. SECURITY MEASURES

Other strategies to improve the robustness of CAPTCHA systems are discussed in this section. Most of them try to highlight the gap between human and machine abilities in solving problems and recognizing objects. These strategies include: - Using sentences as the source of the CAPTCHA instead of characters, words or images: this technique uses human ability in recognizing natural sentences and detecting machine translated mistakes . - Asking users to answer a logical question that requires human thinking. - Using 3D images or characters based on the fact that humans are better than machines in recognizing 3D objects . - Using structures that humans recognize better than machines; such as trees . - Combining text and graphics in the tests: most recognition tools are domain specific; they work exclusively with graphics or text. This strategy can confuse those tools . - Requiring more human interaction, example: using mouse click, dropdown list or drag-n-drop for answering CAPTCHA tests will reduce the risk of blind guessing attacks. CAPTCHAs based on linguistic knowledge: Some current CAPTCHA systems co-ordinate an OCR problem with linguistic knowledge in order to strengthen their tests. Examples of such CAPTCHAs include semCAPTCHA, odd-words-out, number-puzzle-text CAPTCHA, SSCAPTCHA and text-domain CAPTCHA

VIII. ARCHITECTURE DIAGRAM



Functions

- Encryption
- Image/Text Processing
- Base 64 Coder
- Configuration

IX. ALGORITHM

The algorithm wont to create the CAPTCHA doesn't got to be made public, though it's going to be covered by a patent. Although publication can help demonstrate that breaking it requires the solution to a difficult problem in the field of artificial intelligence, deliberate withholding of the algorithm can increase the integrity of a limited set of systems, as in the practice of security through obscurity. The most important factor in deciding whether an algorithm should be made open or restricted is the size of the system. this paper proposed a

replacement recognition algorithm supported holistic verification During the method of this algorithm, Recurrent Neural Network (RNN) was first wont to recognize unknown CAPTCHAs. Then, recognition results were verified by SVM rejection, synthetic data generation and Extreme Learning Machine (ELM). Experiments results show that this algorithm can't only recognize closely-connected CAPTCHAs but also effectively boost the popularity rate of RNN.

X. IMPLEMENTATION

Import classes, load ui (user interface) and display captcha Here, The generated CAPTCHA image is displayed to user with the help of java program and the user gives input accordingly.
In this stage where we get the user input, user's IP address and other required parameters are passed to CAPTCHA server, where the user inputs are validated and returns true or false.
1.Generate a random string and save it to user's session.
2.Write a servlet to generate the CAPTCHA image by using the string.
3.Invoke the servlet to generate the CAPTCHA image and make the user to type it.
4.Check the match between the user's input and session CAPTCHA value.
If the given values match, allow the form submission to continue. If the values doesn't match, then their input is ignored and display an error telling them their input was invalid.

XI.CONCLUSION:

CAPTCHA shows a vital starring role in the life of safety wherever it precludes the Bot hackers from abusing the web service area. This research methodology proposes the CAPTCHA on cloud and provides the scheme to be safe and less defenseless

XII.REFERENCE:

[1]L. Von Ahn, M. Blum, and J. Langford, "Telling humans and computers apart automatically," Communications of the ACM, vol. 47, pp. 56-60, 2004.

[2] A. L. Coates, H. S. Baird, and R. J. Fateman, "PessimalPrint: a reverse Turing test," International Journal on Document Analysis and Recognition, vol. 5, pp. 158-163, 2003.

[3] J. Yan, "Bot, cyborg and automated turing test," in Security Protocols Workshop, 2006, pp. 190-197

[4] H. Baird and K. Popat, "Human interactive proofs and document image analysis," presented at the The 5th IAPR International Workshop on Document Analysis Systems (DAS 2002), 2002.

[5] E. Bursztein, S. Bethard, C. Fabry, J. C. Mitchell, and D. Jurafsky, "How good are humans at solving CAPTCHAs? a large scale evaluation," in 2010 IEEE Symposium on Security and Privacy (SP), 2010, pp. 399-413.

[6] Niket Kumar Choudhary et al, "captchas based on the Principle- Hard to Separate Text from Background" / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6), 2014, 7501-750

[7] Anju Bala and Baljit Singh Saini, "A Review of Bot Protection using CAPTCHA for Web Security", (IOSR-JCE) IOSR Journal of Computer Engine.

[8] S. Ashok Kumar, N. Ram Kumar, S. Prakash and K Sangeetha, "Gamification of Internet Security by Next Generation captchas", International Conference on Computer Communication, IEEE 2017.