iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** II **Month of publication:** February 2022

DOI: https://doi.org/10.22214/ijraset.2022.40190

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue II Feb 2022- Available at www.ijraset.com

A Review on Evolution of various CAPTCHA in the field of Web Security

Mohammad Umar[1], Shaheen Ayyub[2]
[1, 2]*Computer Science & Engineering, Technocrats Institute of Technology*, Bhopal, Madhya Pradesh, India

Abstract: ***CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) which is a security challenge through which human and bots intervention can be measured. It is a type of turing test through which programmed intervention can be detected by it behavior or solving the problem. There are various CAPTCHA problems available such as distorted string, picture recognition, audio, math and gaming CAPTCHA. Game based problem is interactive and highly secured as compare to the other CAPTCHA. In this kind of CAPTCHA user has to solve an AI problem either by drag & drop method or click based, depending on the game. The paper is intended to review various implemented CAPTCHA and compare their weakness and security parameters. Many of the CAPTCHAs are based on click based methods where user has to identify the pictures as per its appearance and click accordingly. But this kind of CAPTCHA can be intervened by image processing techniques such as object classifier. Dragging an object to the target area is an effective way but it has to be performed or solved by an intellectual problem. If dragging an object to the target area by object recognition then system may get cracked by relay attacks.***

Keywords: CAPTCHA, Web Security, Game, Picture Recognition, Math CATPCHA, Image Processing, Relay Attack.

I. INTRODUCTION

Basically, the CAPTCHA was created in the mid 2000s as an approach to telling whether somebody was a human or robot - a kind of Turing Test. The test wasn't totally computerized - humans needed to endeavor to interpret some distorted text - unintelligible to computers - and trust we hit the nail on the head. It did the work. Also with so many web clients finishing these tests consistently, Google saw a chance for something else. In the wake of buying CAPTCHA in 2009 it became reCAPTCHA and we were given something to do translating old bits of writing, whether or not we understood it. Tragically, the free record administration wasn't to endure. A recent report by Google observed that AI robots had the option to translate the CAPTCHAs with 99.8% precision, and numbers in pictures with 90%. Another technique for separating must be found. Albeit this case might look straightforward, there is an extremely refined cycle behind it. Google's investigation works away behind the scenes running its own Turing Test in light of how the client is acting all through their connections on the site. Notwithstanding making it considerably more straightforward for us to finish confirmation processes, designers are continually searching for approaches to making it smoother. Venture forward "The Honeypot" strategy. The Honeypot strategy makes things simpler for clients, while giving a successful technique for getting those troublesome spambots. It has been additionally realized that humans will finish up any problem, as long from their perspective. So imagine a scenario where we made a few imperceptible fields that must be filled in by spambots [1].

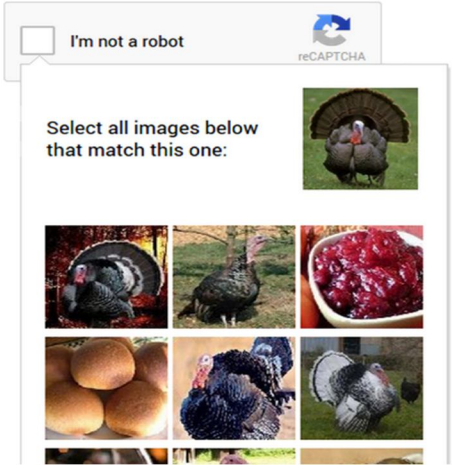


Fig. 1.Now Google's Risk Analysis Engine [1]

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue II Feb 2022- Available at www.ijraset.com

By making the check cycle undetectable, humans aren't pestered by it by any means, and you can feel consoled that those spambots are willing surrendering themselves - particularly when joined with Google's noteworthy examination. With a more modern spam catcher comes more intricate turn of events. There are a few extraordinary instructional exercises web based enumerating how to set it up, so it merits putting resources into. The vital thing to keep an eye out for is guaranteeing your clients can in any case utilize autocomplete, without being hailed as a robot [1]. There are various scenario where risk analysis can be measured by facing certain hard AI based problems but problem should be easier for human and harder for robot and can be solved within a second.

II. RELATED WORKS

A. Related Works

Philip Kirkbride et al. [2] proposed a CAPTCHA for intrusion detection. In this paper, creators considered the clever utilization of game-like CAPTCHA as a strategy for information assortment to be utilized in making a conduct biometric for recognizing deceitful record use. Given the requirement for adjusting security without compromising client experience, it is accepted that game-like CAPTCHA for the age of biometric social information can give an answer. Creators suggest that a model game-like CAPTCHA be made and carried out as a feature of an interruption identification frameworks (IDS) for additional review on the adequacy of game-like CAPTCHA for account confirmation. On the off chance that such a game is made utilizing front-end web innovation like JavaScript and HTML, it will be generally simple to catch the information utilizing a previous library like rrweb.io. After each game-play the information gathered will be shipped off a server side data set. To recreate genuine utilization, numerous guineas pigs will be welcome to utilize the CAPTCHA on various days, rather than a few times in succession. The underlying 5-10 game plays will be utilized as the enlistment time frame and meetings after that will be utilized with a SVM calculation to decide whether it can precisely separate the first player from others. Accepting exact outcomes from this one-class SVM calculation, creators might endeavor to additionally work on the pace of ID by joining other client credits like IP, client specialist, time-region, and login-time. Monther Aldwairi et al. [3] proposed a system that assessed another sort of CAPTCHA: Flash-based CAPTCHA. As per the overview results, this CAPTCHA was the most advantageous to use since it was casted a ballot the simplest to settle, with the most un-number of disappointments. Also, it was viewed as the quickest, the most lovely to tackle, and the least demanding to recollect after not involving it for a significant length of time. Additionally, Flash-based CAPTCHA needs less assets contrasted with the current CAPTCHAs, making it more proficient for use. Streak based CAPTCHA is impervious to OCR attacks since this attack targets text-based CAPTCHAs, and the way that this CAPTCHA needs mental capacities to address, which makes it more impervious to computerized attacks. Also, clients from various age gatherings, levels of training, Internet abilities and surprisingly those with vision weaknesses had the option to address it without any problem.

a. Flash-Based CAPTCHA

b. correctly solved CAPTCHA

Fig. 2.Drag and Drop Based Games [3]

Ahmet Faruk Çakmak et al. [5] proposed an audio CAPTCHA which is based on RastaPLP Features by SVM. The Naïve Bayes strategy accurately distinguishes around 42% of the test digits. This strategy likewise fizzled in light of the fact that each class component in the train set isn't adjusted in light of the fact that the train set has an enormous number of commotion class (eleventh class) components, while the components from 0 to 9 are fairly less. None of the 100 sound documents in the test set were completely perceived by Naïve Bayes strategy. Since the quantity of components isn't adjusted, despite the fact that the classes from 0 to 9 to some degree perceived even in the train set, the commotion class has a generally low achievement pace of 71%. To this end all test sound documents are not perceived without blunder, yet regardless of whether a part is wrongly doled out to the clamor class, it implies that the test component is misclassified. Nitisha Payal et al. [6] proposed a CAPTCHA which is based on hybrid images. AJigJax is a drag-drop based Captcha in a type of straight jigsaw puzzle. The proposed work presents two levels in Captcha, one is CL1: AJigJax; for those sites that are seldom gotten to or need less security or no validation and other one is CL2: AJigJax; for those sites that have basic data and need confirmation to be done and are regularly gotten to. Based on the exhibition assessment of Captcha, we can emphatically say that AJigJax Captcha is effectively addressed, engaging, less tedious, easy to understand. AJigJax is safer as simplified are gotten than composing text to pass the test. CL2: AJigJax must be addressed by legitimated client as the idea of graphical secret key is added to it.

Cao Lei et al. [7] proposed a CAPTCHA which is based on finger guessing game makes machines to make a second logic judgment on the basis for the identification, improved the difficulty for machines to pass. The finger-guessing game has the broad foundation of the population, so the CAPTCHA obviously reduces the difficulty of human recognition. It is a progress of the existing image verification code technology field. But finger guessing game is not an intellectual approach through which is server can be secured more precisely. Sometimes finger guessing game can become more confusing because of various gestures prompting in the screen for recognition that degraded the performance of the CAPTCHA and may irritate users to interact with it.

Fig. 3.All Finger Gestures [7]

Hong Yu et al. [8] proposed an Automatic Generation of Game-based CAPTCHA. In the fundamental execution, the game based CAPTCHA utilizes text based idea marks. Along these lines a bot furnished with PC vision abilities can undoubtedly perceive the text in the game. Yet, to break the CAPTCHA, the bot likewise needs to reason about the connection between the ideas, either through looking through on the web or breaking into the information data set. Despite the fact that we extricate the underlying information data set from ConceptNet which is freely available, the AGCG framework can be handily conveyed with a private information data set which is safer for business use. In a perfect world, private information data set has relations that don't cover altogether with public rational information data sets because of the inadequacy of information data set and the enormous measure of conceivable conventional relations. It might take a piece longer for players to complete proposed game based CAPTCHA than a conventional visual put together CAPTCHA with respect to a personal computer. In any case, the games might carry more delight to a client than an OCR task, not entirely settled. Proposed game-based CAPTCHA could be more proper for portable conditions where it is more straightforward for the clients to swipe and haul than to type in words. The CAPTCHA is a significant instrument to keep bots from getting to web administrations. A developing examination local area is concentrating on the best way to construct new CAPTCHAs that are impervious to bots while simple for humans. Proposed naturally created game-based CAPTCHAs join the security of the conventional visual based CAPTCHAs, the human agreeableness of the rationale based CAPTCHAs, and the fun of PC games. It can naturally create enormous enough number of game based CAPTCHAs to forestall straightforward savage power attacks. Hence we accept that proposed game-based CAPTCHAs are equipped for establishing a safer climate on Internet and giving a superior web administration to clients.

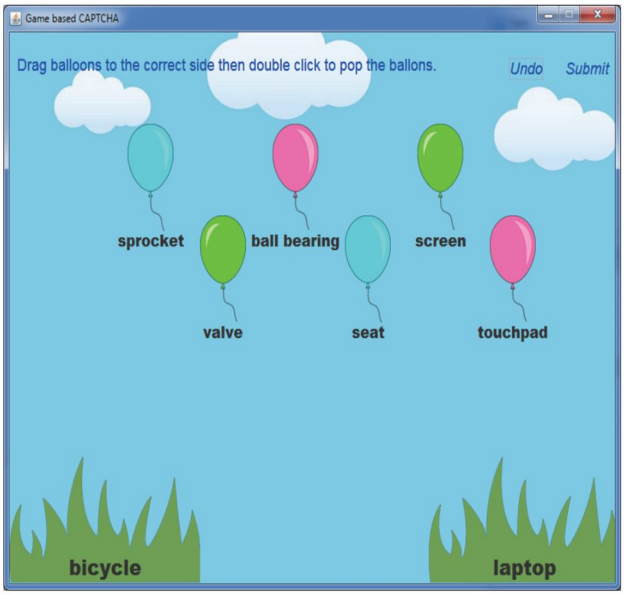


Fig. 4. A screenshot of a preliminary game-based CAPTCHA [8]

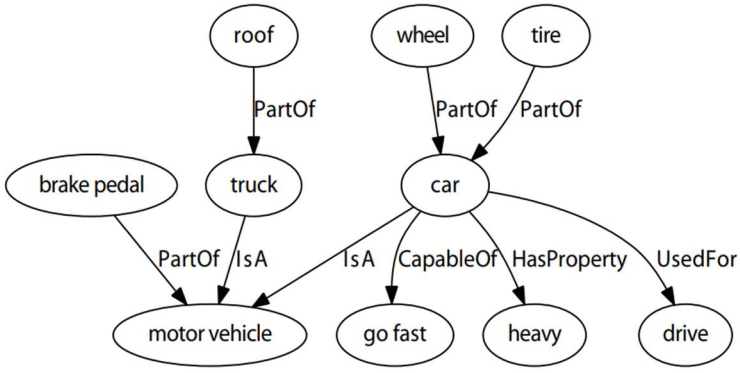


Fig. 5. A subgraph extracted from the Concept-Net [8]

Shardul Vikram et al. [9] proposed non-intrusive moving-target defense system named as NOMAD. NOMAD keeps web bots from mechanizing web asset access by randomizing HTML components while not influencing typical clients. In particular, to forestall web bots remarkably distinguishing HTML components for later mechanization, NOMAD randomizes name/id boundary upsides of HTML components in every HTTP structure page. As per the assessment, NOMAD can forestall this large number of web bots with a generally low upward. NOMAD can be normally carried out at the server-side by adjusting the source code of the web applications. Additionally, NOMAD could be executed as middleware between the server and customer, to try not to add the intricacy to the server side rationale of the web applications. Carrying out NOMAD as a middleware permits it to be free and all around appli-link to various web applications (without straightforwardly changing the source code) and customer side technologies (e.g., various programs and modules). Along these lines, the middleware arrangement will be straightforward to the two servers and end clients. Zhen Li et al. [10] proposed a CAPTCHA which is based on game theory. In this paper, we formally modeled the interdependence of the decision-making by the defender and the attacker in a Stackelberg game theoretic framework. Through best response and strategy analysis, the break even points of whether adopting machine solver or human solver can be determined. In contrary to traditional wisdom to make CAPTCHA harder, we proposed two models that feature easy CAPTCHA with time latency constraints as well as incorporation of cryptocurrency mining into existing CAPTCHA mechanism. The results discourage attackers from using human solvers and generate a welfare-enhancing CAPTCHA business model. Aadhirai et al. [11] proposed a system which is based on vision where user will have to identify the object based on distance. Proposed system serves an image of real world where different kind of objects relies. System raises an artificial problem where user will have to recognize a particular object which is farthest from a specified object. It may difficult to recognize for those person who has poor vision because there is a hazy appearance which is difficult for normal human also. If it is possible to observe then it can be only done by human not by bots. It is highly secured CAPTCHA which having difficult artificial problem which is impossible to solve by bots. Ibrahim et al. [12] proposed a system in which user will have to rotate the cube and identify the respective colors whereas marked with question marks. Once the user is able to rotate and identified the character mentioned over 3D cube, system allow user to get accessed otherwise a new problem will be served and color model will get changed and a new challenge proposes. Text box and 3D cube both have identical colors and user requires to match both the colors and recognize the correct letter and type over there for successful turing test.

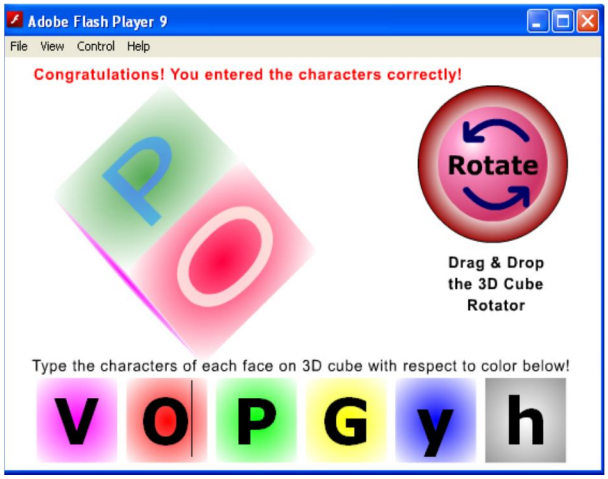


Fig. 6.3D Cubic CAPTCHA [12]

III. PROBLEM IDENTIFICATION

S. Ezhilarasi et al. [4] proposed a system which is based on image recognition named as IRA (Image Recognition Annotation). In this system, authors distorted the image by resizing, morphology and transparency. System added some noise in the images and makes it complicated for bots to be processed. But some time adding noise in the image makes it complicated for human also. CAPTCHA should be as easy as possible for human and should not take too much time. It means that CAPTCAH should be easy, less time consuming, less space complexity and high secured. Now gaming CAPTCHA is in trend and requires certain attention from user by making it interesting. But not image processing based approaches like google lens that works with tensorflow and yolo based techniques are much more efficient to recognize and classify the objects from images that can crack the security premises of the image recognition based CAPTCHA.

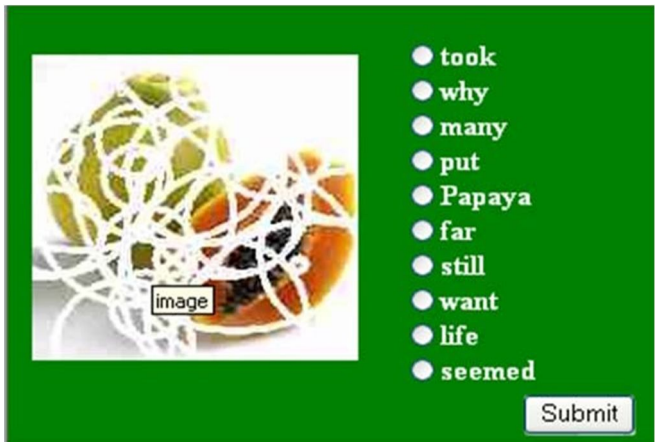


Fig. 7.IRA CAPTCHA for Distorted Picture [4]

Fig. 7 shows the IRA CAPTCHA where picture has been distorted and user is required to identify the piture and click on radio button accordingly. But sometime distortion level turns it more complicated for human too that may irritate users.

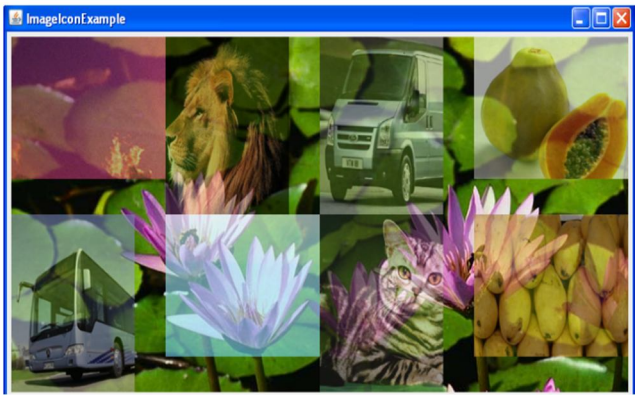


Fig. 8.IRA CAPTCHA for Overlapped Picture [4]

Fig. 8 shows the IRA CAPTCHA where pictures have been overlapped with other pictures and user is required to identify the picture and click accordingly. It may confuses the users to get the actual one.

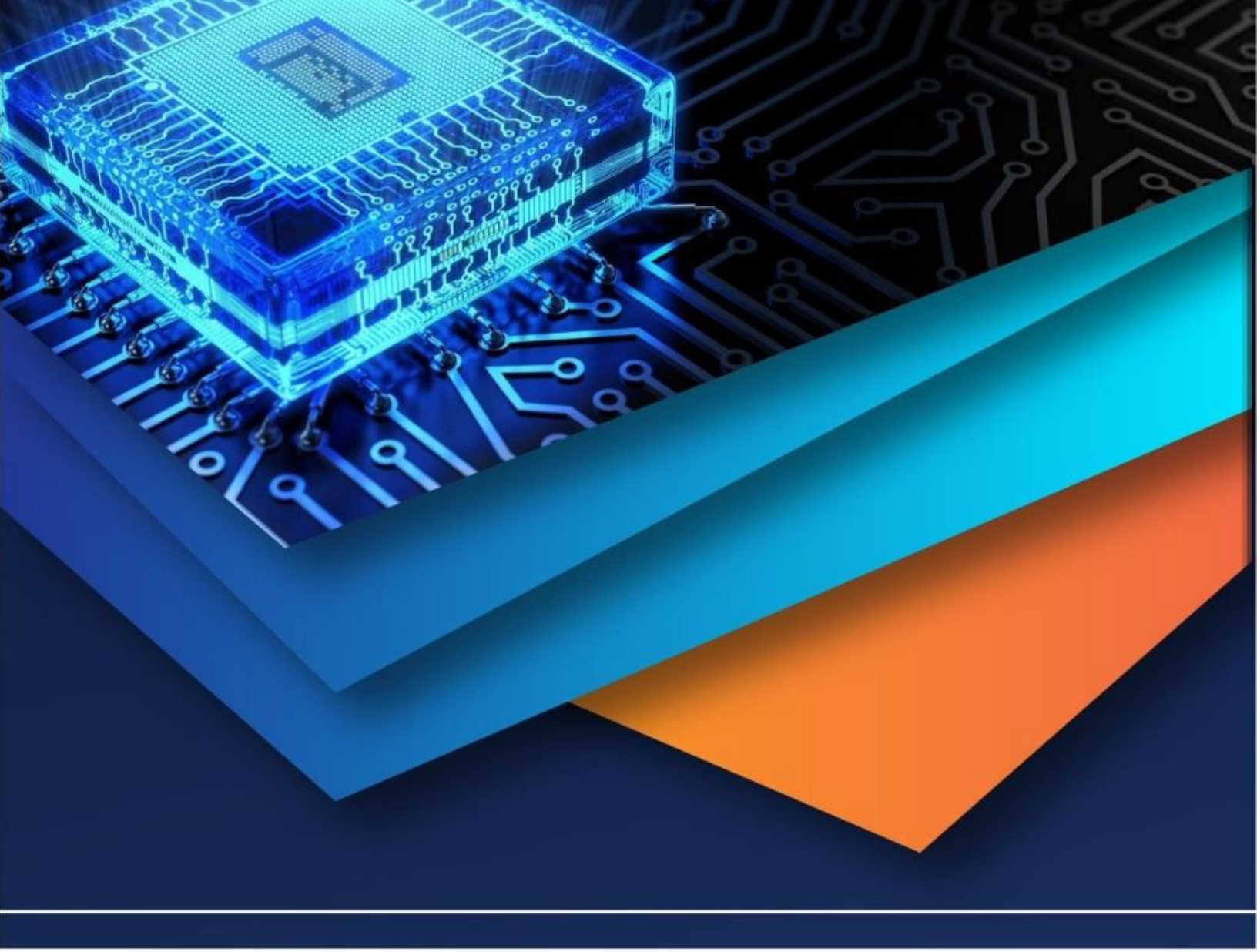
IV. CONCLUSION & FUTURE SCOPE

The intension of the paper is to review various implemented systems in the field of CAPTCHA. Most of the systems have been used picture recognition CAPTCHA where pictures may be in original appearance or distorted form. Normal picture can be recognized using machine learning approaches and distorted one get confuses human too. Certain systems are based on flash gaming but game level is bit lower and often easy for bot too. Dragging an object to the target position is not an intellectual approach. A gaming CAPTCHA now can be enhanced and become more intellectual to secure the web premises more accurately. Game may be decision based or it can be stated as decisive games. Decisive game can be often easy for human but almost impossible for robots.

REFERENCES

[1] Adapt, CAPTCHA, 2018. [Online]. Available: https://www.adaptworldwide.com/insights/2018/the-evolution-of-captcha, [Accessed: 29- Jan- 2022]
[2] P. Kirkbride, M. A. Akber Dewan and F. Lin, "Game-Like Captchas for Intrusion Detection," IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress, 2020, pp. 312-315.
[3] Aldwairi, Monther & Mohammed, Suaad & Padmanabhan, Megana. (2020). Efficient and Secure Flash-based Gaming CAPTCH.
[4] S. Ezhilarasi and P. U. Maheswari, "Image Recognition and Annotation based Decision Making of CAPTCHAs for Human Interpretation," 2020 International Conference on Innovative Trends in Information Technology (ICITIIT), 2020, pp. 1-6, doi: 10.1109/ICITIIT49094.2020.9071558.
[5] Cakmak, Ahmet & Balcilar, Muhammet. (2019). Audio Captcha Recognition Using RastaPLP Features by SVM.
[6] N. Payal and R. K. Challa, "AJIGJAX: A hybrid image based model for Captcha/CaRP," 2016 IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics Engineering (UPCON), 2016, pp. 38-43, doi: 10.1109/UPCON.2016.7894621.

[7] Cao Lei, "Image CAPTCHA technology research based on the mechanism of finger-guessing game," Third International Conference on Cyberspace Technology (CCT 2015), 2015, pp. 1-4, doi: 10.1049/cp.2015.0843.
[8] Yu, Hong and Mark O. Riedl. "Automatic Generation of Game-based CAPTCHAs." (2015).
[9] S. Vikram, Chao Yang and Guofei Gu, "NOMAD: Towards non-intrusive moving-target defense against web bots," 2013 IEEE Conference on Communications and Network Security (CNS), 2013, pp. 55-63, doi: 10.1109/CNS.2013.6682692.
[10] Z. Li and Q. Liao, "CAPTCHA: Machine or Human Solvers? A Game-Theoretical Analysis," 2018 5th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2018 4th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom), 2018, pp. 18-23, doi: 10.1109/CSCloud/EdgeCom.2018.00013.
[11] Aadhirai R, Sathish Kumar P J and Vishnupriya S, "Image CAPTCHA: Based on Human Understanding of Real World Distances" Proceedings of 4th International Conference on Intelligent Human Computer Interaction, IEEE 2012.
[12] Ibrahim FurkanInce, YucelBatu Salman, Mustafa ErenYildirim and Tae-Cheon Yang, "Execution Time Prediction For 3D Interactive CAPTCHA By Keystroke Level Model" in Fourth International Conference on Computer Sciences and Convergence Information Technology of IEEE 2009.
[13] JingSong Cui, LiJing Wang, JingTing Mei, Da Zhang, Xia Wang, Yang Peng, WuZhou Zhang, "CAPTCHA Design Based on Moving Object Recognition Problem" in IEEE 2009.
[14] Jing-Song Cui, Jing-Ting Mei, Xia Wang, Da Zhang, Wu-Zhou Zhang , "A CAPTCHA Implementation Based on 3D Animation" in International Conference on Multimedia Information Networking and Security of IEEE 2009.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY