International Journal of Interdisciplinary Innovative Research &Development (IJIIRD)
ISSN: 2456-236X
Vol. 05 Issue 02 | 2021

A Review on the Security and Convenience of CAPTCHA Schemes

Anjana Menon R[1], Feon Jaison[2]

[1]*Master student, Department of MCA, Jain Deemed-to-be University, Bengaluru, Karnataka, India*
[2]*Assistant Professor, Department of MCA, Jain Deemed-to-be University, Bengaluru, Karnataka, India*

ABSTRACT

Billions of people are using Internet at every moment. Internet brings security issues along with the content we access. CAPTCHA - Completely Automated Public Turing test to tell Computers and Humans Apart is a common and widely used technique that is used in websites for securing the web content. The basic function of CAPTCHA is to differentiate between a legitimate human user and malicious bots. The paper is a comparative study of two different CAPTCHA schemes in terms of security and convenience. Here, the CAPTCHA schemes studied are ReCAPTCHA by Google and two layer CAPTCHA from Microsoft. Both of these CAPTCHA schemes are discussed in detail in terms of the security it provides. In most of the cases, a highly secured CAPTCHA scheme is difficult for human users to solve. This causes usability issues and finally it will last in a situation where the users decide to leave the site. The main motive of this work is to discuss and identify a CAPTCHA scheme that is not compromising security over convenience and vice versa.

Keyword: - CAPTCHA, ReCAPTCHA, Security, Convenience.

1. INTRODUCTION

In this digital era, Internet is used excessively. Here comes the importance of web security. For this, several security measures like firewalls, anti-virus, IDS, IPS etc. can be used but from the end user perspective, CAPTCHA challenge is of great importance. CAPTCHA is using a challenge-response strategy which is actually like a small question. This question should be correctly answered by the user to authenticate themselves as the true owner or legitimate data user. A CAPTCHA may seem meaningless or irritating but it actually does a lot to secure the web content. The CAPTCHA technique authenticates whether the requesting user is a human or a malicious bot. This can block spamming and automated malicious requests. Bots have the ability to sign up in several websites and this may cause harm to those websites. All these kinds of abuses can be prevented using CAPTCHAs. The uses of CAPTCHA includes prevention of junk mails, protection against unauthorized access, avoiding automated polling, prevention of comment spamming and ensuring secure online transactions.

The common CAPTCHA schemes are divided into three categories: Text based, Image based and Audio based. Text based CAPTCHAs include English alphabets and Arabic numerals. A difficult CAPTCHA can be created by adding noise or by using sophisticated distortion. A number of CAPTCHAs schemes are being introduced since its evolution. Some of the mentions are: math problem, word problem, social media log in, time based CAPTCHA, honeypot CAPTCHA, picture identification challenge, no CAPTCHA ReCAPTCHA, invisible ReCAPTCHA, confident ReCAPTCHA, sweet CAPTCHA and so on. All of these updated CAPTCHA schemes are meant to secure the websites more but many of the schemes proved inconvenient. It is always important for a CAPTCHA scheme to be secure and usable.

In this paper, Microsoft two layer CAPTCHA and Google ReCAPTCHA are being discussed and compared in terms of security and convenience. Section A explains about the two layer CAPTCHA scheme and section B discusses about the ReCAPTCHA concept. Section C is showing the comparison study between the two CAPTCHA schemes. In section D, the paper is concluded by focusing on the CAPTCHA scheme that is more secure and convenient.

2. TWO LAYER CAPTCHA

Two-layer CAPTCHA is a text CAPTCHA scheme introduced by Microsoft in 2015. This is the very first CAPTCHA scheme that that uses two layers of text as the CAPTCHA challenge. Most traditional text CAPTCHAs are single layered. Microsoft two layer CAPTCHA is a combination of two single layered CAPTCHAs. Two layer CAPTCHA is an image with six characters arranged in two layers. Here, both layers are having three characters each. This kind of CAPTCHA challenge includes both solid and hollow characters. The hollow characters are created using contour lines.

The generation of this CAPTCHA challenge includes a number of steps:

Step 1: A set of six characters are chosen randomly from the entire character set. Secondly, Step 2: Some random characters are selected from those six characters and convert them to hollow characters.

Step 3: Combine all the six characters to form a CAPTCHA with three characters per line.

Step 4: The character combination is warped and rotated to form the final two layer CAPTCHA challenge.

The response to the challenge should be made from the top left character. The characters from the first horizontal layer is entered first. For a successful CAPTCHA response verification, the users have to enter the characters layer by layer without any space in between.

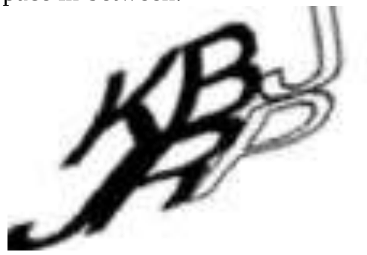


Fig: 1 Microsoft Two Layer CAPTCHA

The above figure [Fig: 1] is an example for Microsoft two layer CAPTCHA. The CAPTCHA includes the characters: ‘K’, ‘B’, ‘J’ in the first layer and ‘J’, ‘R’, ‘P’ in the second layer. The characters ‘J’ and ‘P’ at the end of each layer are the hollow characters created randomly.

3. ReCAPTCHA

ReCAPTCHA is invented in 2007. It is a free service provided by Google for web protection. ReCAPTCHA is one of the most widely used CAPTCHA scheme. It has been welcomed by many popular websites for preventing automated malicious bots from conducting nefarious activities. ReCAPTCHA is currently being used by more than six million websites. Current ReCAPTCHA versions include: ReCAPTCHA V2, ReCAPTCHA V3 and ReCAPTCHA Enterprise.

ReCAPTCHA V2 was introduced in 2014 which is based on an *advanced risk analysis system that* relies quite heavily on Google cookies. ReCAPTCHA V2 challenges include a check box that says *“I’m not a robot”* and also will challenge you with an image or audio recognition task. Chrome users commonly get a check box challenge, while firefox users get a difficult image recognition challenge. The users will get tougher challenges based on their behavior in a website, which will degrade their user experience.

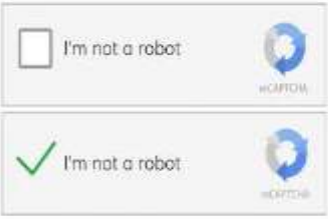
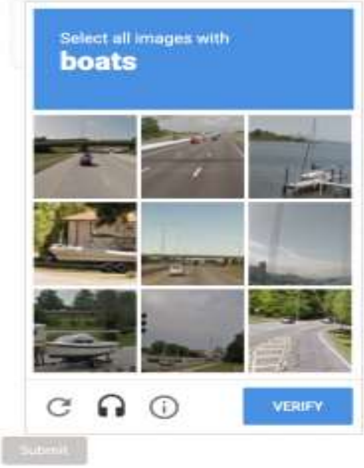


Fig 2: ReCAPTCHA V2 – Check box challenge

The above image (Fig 2) is showing the check box challenge. The first one is the challenge and the second one (checked box) is the response.



Select all images with
boats

Fig 3: ReCAPTCHA V2 – Image Identification Challenge

The above figure [Fig 3] is showing the image identification challenge where the users have to identify all those images that shows boats.

ReCAPTCHA V3 is an updated version of V2 that provides a better user experience. Unlike V2, V3 is having no challenges. Instead of creating challenges, ReCAPTCHA observes user behavior to identify whether the user is a human or bot. With every user request the, reCaptcha v3 assigns a score between 0 and 1. If the score is close to 0: it tells 'sorry, you're a bot' or if it is close to 1: it tells 'congrats, you're a human'. V3 improves the usability as it is not disrupting the users with ReCAPTCHA challenges.

Fig 4: ReCAPTCHA V3

ReCAPTCHA Enterprise is similar to ReCAPTCHA V3 where the users can enjoy seamless website experience. Here, the users are not given any CAPTCHA challenge. ReCAPTCHA Enterprise can be used to detect automated attacks or threats against the website that originate from scripts, mobile emulators, bot software, or humans.

Fig 5: ReCAPTCHA Enterprise

Table I: Comparison of ReCAPTCHA Versions

Features	ReCAPTCHA Enterprise	ReCAPTCHA V3	ReCAPTCHA V2
Proven	Yes	Yes	Yes
Customer Friendly	Yes	Yes	No
Customizable	Yes	No	No
Adaptive	Yes	No	No
Comprehensive	Yes	No	No
Enterprise Ready	Yes	No	No
Customer Lead	Yes	No	No
High Precision Risk Scores	Yes	No	No

(Results of comparison study)

4. COMPARISON STUDY AND RESULTS

The paper has discussed the two CAPTCHA techniques in detail. Both Microsoft two layer CAPTCHA and Google ReCAPTCHA are having their own importance in securing web content. Microsoft Two Layer CAPTCHA is a secure CAPTCHA technique but it proves difficult at times. This kind of CAPTCHA is prone to certain attacks like segmentation and recognition attacks. This highly distorted CAPTCHA is difficult for humans as well. ReCAPTCHA is comparatively more secure and convenient. Among the three versions, Enterprise version is the most secured one. ReCAPTCHA Enterprise is the CAPTCHA scheme that is equally secure and convenient.

Table II. Comparison Study of Two Layer CAPTCHA and ReCAPTCHA

	2 Layer CAPTCHA	ReCAPTCHA
Security	Low	High
Convenience	Low	High

(Results of the comparison study)

5. CONCLUSION

The paper has systematically analysed the security and convenience factors of a two layer CAPTCHA deployed by Microsoft and ReCAPTCHA technology developed by Google. While designing a CAPTCHA, we are concerned about the level of security it offers for the website protection. For providing security, a CAPTCHA challenge should always be difficult for bots. By creating difficult CAPTCHAs, sometimes it becomes impossible for human users to solve the challenge. Hence, a CAPTCHA is worth implementing if it is secure and convenient. The paper has studied and identified Google ReCAPTCHA as the most secured and convenient CAPTCHA scheme among the other popular and currently used ones. This evolving and revolutionary technology possess great importance in web security and will be continued by millions of websites.

6. REFERENCES

[1] Rachana.B.S, Dhruthi.S, Swarna.R, Chandan.A, IMPROVED SECURITY ASPECTS ON MICROSOFTS TWO -LAYER CAPTCHA, IJARIIE-ISSN(O)-2395-4396, Vol-2 Issue-5 2017.

[2] Gao, Haichang & Tang, Mengyun & Liu, Yi & Zhang, Ping & Liu, Xiyang. (2017). Research on the Security of Microsoft's Two-Layer Captcha. IEEE Transactions on Information Forensics and Security. PP. 1-1. 10.1109/TIFS.2017.2682704.

[3] Ahn, Luis & Maurer, Benjamin & McMillen, Colin & Abraham, David & Blum, Manuel. (2008). reCAPTCHA: Human-based character recognition via Web security measures. Science (New York, N.Y.). 321. 1465-8. 10.1126/science.1160379.

[4] Gafni, Ruti & Nagar, Idan. (2016). CAPTCHA – Security affecting User Experience. 905. 10.28945/3469.

[5] Lung, Jonathan. (2012). Ethical and legal considerations of reCAPTCHA. 2012 10th Annual International Conference on Privacy, Security and Trust, PST 2012. 211-216. 10.1109/PST.2012.6297942.

[6] Gao, Haichang & Yan, Jeff & Cao, Fang & Zhang, Zhengya & Lei, Lei & Tang, Mengyun & Zhang, Ping & Zhou, Xin & Wang, Xuqin & Li, Jiawei. (2016). A Simple Generic Attack on Text Captchas. 10.14722/ndss.2016.23154.

[7] J. Yan and A. S. El Ahmad, "Captcha Robustness: A Security Engineering Perspective," in Computer, vol. 44, no. 2, pp. 54-60, Feb. 2011, doi: 10.1109/MC.2010.275.

[8] Pengpeng Lu, Liang Shan, Jun Li and Xunwei Liu, "A new segmentation method for connected characters in CAPTCHA," 2015 International Conference on Control, Automation and Information Sciences (ICCAIS), 2015, pp. 128-131, doi: 10.1109/ICCAIS.2015.7338647.

[9] Y. Zi, H. Gao, Z. Cheng and Y. Liu, "An End-to-End Attack on Text CAPTCHAs," in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 753-766,2020, doi:10.1109/TIFS.2019.292862.

[10] Chen, Jun & Luo, Xiangyang & Guo, Yanqing & Zhang, Yi & Gong, Daofu. (2017). A Survey on Breaking Technique of Text-Based CAPTCHA. Security and Communication Networks. 2017. 1-15. 10.1155/2017/6898617.

[11] Banday, M. Tariq & Shah, N. (2011). A Study of CAPTCHAs for Securing Web Services.