International Journal of Latest Trends in Engineering and Technology
Special Issue April-2018, pp. 005-011
e-ISSN:2278-621X

A STUDY ON CAPTCHAs THE CHALLENGE RESPONSE TEST

A.Abiya jecinth kumar[1], A.R.Guru Gokul[2]

Abstract –In the world of humongous security threats it has became mandatory for a website to use a CAPTCHAs where CAPTCHAS plays a vital role in the web security by differentiating bots and humans. It acts as a security barrier that identifies the bots and stops its access. CAPTCHAS stands for "Completely Automated Public Turing test to tell Computers and Humans Apart". This paper covers the different types of CAPTCHAS, Attacks on CAPTCHAS, hacking of CAPTCHAs and some future CAPTCHAs that include the regional languages and gaming ideas.
Keywords –CAPTCHAs, Security, Attacks,Hacking , Future CAPTCHAs, regional languages, games.

1.INTRODUCTION

Many People may have experienced a security question while accessing any websites such as Gmail, twitter, yahoo, Face book; etc .out of those many would have been for the login purpose this security question is called as CAPTCHAs. They're also known as a type of Human Interaction Proof (HIP). CAPTCHAS is the type of challenge-response test used in computing to ensure that the response is not generated by the bots .The process usually involves one computer asking a user to complete a simple test .The test may consist of letters, digits, audio or images that any user entering a correct solution is accepted as a human and the user who fails to enter the correct solution is considered as a Robot and the access to the desired page is denied. This CAPTCHAs initially started with the text and digits and later the audio CAPTCHAs was introduced and the latest CAPTCHAs used is the images. The aim is to create a test that humans can pass but not the computer. It's also important that the CAPTCHAS application is able to present different CAPTCHAs to various users at each time. If a visual CAPTCHAS presents a static image that was the same for every user, the spammer can easily recognize, decode the letters, and develop an application to type the correct answer automatically. All the CAPTCHAS don't rely on visual patterns. It's important to have an alternative to a visual CAPTCHAS. One alternative to a visual test is an audible test. An audio CAPTCHAS usually presents the user with a series of words, letters or numbers. It's not unusual for the program to distort the speaker's voice, and it's also common for the program to include background noise in the recording. The ordinary CAPTCHAs has been outdated as many deep learning algorithms and pattern recognitions algorithm can decode this CAPTCHAs that makes an urge for the advanced CAPTCHAs. This paper shows the future ideas of the CAPTCHAs that is more safe as only human could solve and keeps the bots away from the access [4].

2.VARIOUS TYPES OF CAPTCHAS

The CAPTCHAs can be basically distinguished as visual and non-visual CAPTCHAs ,apart from this there are many types that are used in a regular CAPTCHAs test .they are as follows

2.1 Text-Based Captchas

These are the most common and simple form of CAPTCHAs where a set of letters or numbers are displayed as a test. To make it little improved twisted and twirled texts were introduced This system, was originally aimed to help digitize printed text that was hard to read by OCR (Optical Character Recognition) software. The other implementation of text based CAPTCHAs are as follows [1].

2.1.1 GIMPY Captchas

Gimpy text CAPTCHAs shown in Figure 1.1 are those that display the text in distorted manner where it makes difficult for the computers to enter the correct answer. The texts are overlapped such that its test the ability for the human himself. This was built by CMU and the yahoo for their messenger services [1].

[1] PG Scholar,Department of IT,Sri Venkateswara College Of Engineering, Sriperumbudur,Tamilnadu,India.
[2] Assistant Professor ,Department of IT,Sri Venkateswara College Of Engineering., Sriperumbudur,Tamilnadu,India.

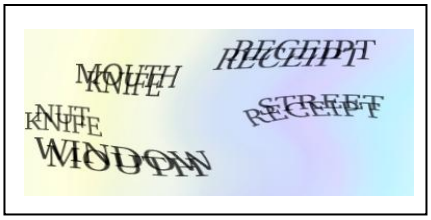

Figure 1.1 –Gimpy CAPTCHAs

2.1.2 Ez-Gimpy Captchas

This is a reduced version of the Gimpy CAPTCHAs shown in Figure 1.2.This selects a word and applies distortion but does not overlap the words and asks the user to find the correct word [1].

Figure 1.2 –Ez-Gimpy CAPTCHAs

2.1.3 Baffle Text Captchas

The Baffle text CAPTCHAs shown in Figure 1.3 picks the alphabets randomly and some half visible distortion is added to that random text .unlike the Gimpy CAPTCHAs it is hard to recognize the word for the computer.

Figure 1.3 –Baffle text CAPTCHAs

2.1.4 Msn Captchas

The MSN CAPTCHAs shown in Figure 1.4 is a type of the CAPTCHAs that uses the 8 letter combo of letter and digit in dark blue colour and with grey background. To make it hard to recognize it makes the text to warp.

Figure 1.4 –MSN CAPTCHAs

2.2 Image Based Captchas

The next level of text CAPTCHAs is the graphical CAPTCHAs that deals with the images the general types of this graphical CAPTCHAs is given below [1].

2.2.1 Pix Captchas

From a given Set of images a group of similar image is expected to select by the user. There is huge collection of images present in the database for the display and out of them a random six to nine images are given as a test this is called PIX CAPTCHAs shown in Figure 1.5.

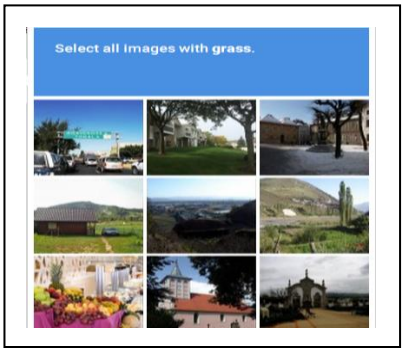

Figure 1.5 –PIX CAPTCHAs

2.2.2 Bongo Captchas

The Bongo CAPTCHAs shown in the Figure 1.6 displays two set of blocks where the user must identify the different symbol that varies in that two blocks.

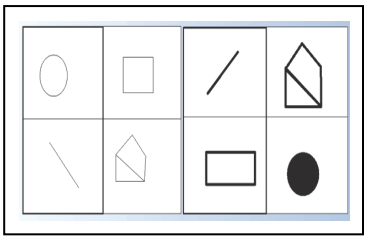

Figure 1.6 –Bongo CAPTCHAs

2.3 Audio Based Captchas

The Audio based CAPTCHAs shown in Figure 1.7 is more advanced than the text based or image based CAPTCHAs as it deals with the strong listening skill of the human .A set letters or number is pronounced and the user is required to listen and type the correct answer to enter the desired site. The main issue with these audio based CAPTCHAs is they need a noiseless atmosphere for the user for the audio clarity [1].

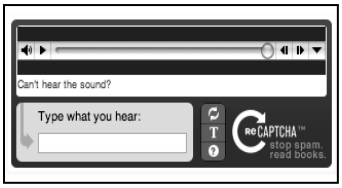

Figure 1.7 –Audio CAPTCHAs

2.4 Maths Based Captchas

The CAPTCHAs that is based on the simple mathemetical functions are shown in Figure 1.8 where a simple problem of DMAS is given to solve.

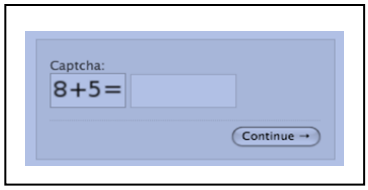

Figure 1.8 –Maths based CAPTCHAs

3. COMMON ATTACKS ON CAPTCHAS

However a CAPTCHAs being more secure there will be always a improved and advanced threats and attacks that could decode the CAPTCHAs by the bots. There are n numbers of attacks that can be categorized as three main categories they are given below in this sub section.

3.1 Basic Captchas-Decode Attacks

This type of attack involves in developing the software or algorithms that could solve and break the CAPTCHAs automatically. It uses the computer vision technique for the automatic solving as most of them are visual based CAPTCHAs and makes the bots to understand the visual easily that initially which was not possible by them.

3.2 Bypassing Attacks Of Captchas
When CAPTCHAs is asked as a test sometimes the hackers would overcome that CAPTCHAs by bypassing the CAPTCHAs instead of solving it. The attacks the skips the CAPTCHAs make the server to think that the CAPTCHAs has been solved and it loads to the site.

3.3 Human Attacks Of Captchas
The CAPTCHAs attacks may not only come from the bots but there is a possible of attacks that could happen by humans too. The humans may do this knowingly or unknowingly, the one who does that knowingly is because of some organization or bots the force to do so. The unknowing attack may occur by bots placing their desired site's CAPTCHAs instead of the original site.

4. CAPTCHAS HACKING IDEAS
When there is a strong barrier in the form of a CAPTCHAs to enter into a certain site, there is a necessity for the hackers to develop a algorithm or develop a certain steps to overcome the CAPTCHAs. In the various types of CAPTCHAs that is described in section 2, the simple form of text based CAPTCHAs is the weakest among all other CAPTCHAs to be easily hacked by the hackers. The other way of easy hacking is when the server repeatedly displays the same CAPTCHAs or displays a certain CAPTCHAs at a regular frequency then it is easy for the hacker to learn the decoding of that CAPTCHAs . Even though the strong CAPTCHAs such as the image based CAPTCHAs is difficult to decode, hacking is possible with the improved algorithms and steps.
It is not easy for a hacker to develop a certain algorithm that solves the CAPTCHAs at a stretch so the hacker develops the algorithm step by step that it makes the job easier for the decoding of the CAPTCHAs. For example they breakdown the CAPTCHAs words in the basis of the pattern, colour, shape of the alphabet, but it makes difficult when distortion is applied. Some of these CAPTCHAs breaking techniques are given below in this subsection [2].

4.1 Hacking In The Absence Of Ocr
Without the optical character recognition (OCR) the CAPTCHAs can be hacked. Sometimes when correct CAPTCHAs is entered it maintains the session without terminating it. With the help of this session id the site can be resumed avoiding the CAPTCHAs test.
This type of hacking comes under the bypassing attack where the server is restored without attempting the CAPTCHAs. Initially connect to the CAPTCHAs page and record the session id and the CAPTCHAs text, then resend tat session id and the text changing the user id. A traditional CAPTCHAs hacking software involves using the image recognition routines to decode the images. This makes it easy to hack the CAPTCHAs images [3].

4.2 Hacking Visual Based Captchas
The most common visual CAPTCHAs is the Gimpy CAPTCHAs where these are generally used in most of the sites. Hackers study the pattern shape and colour of the letters and the numbers that are displayed and develops the algorithm to decode the CAPTCHAs. Consider finding the letters "j","e","w","e","l" and connecting them to form a word "jewel" is similar to finding various parts and connecting them to form a human. Considering the Ez-Gimpy CAPTCHAs there is a three step algorithm that can be hacked. These are listed below [2].

4.2.1 Locate Letters At Various Locations
The initial step is to propose a set of letters in the image using the shape matching techniques. This method matches the image at random and compares the letters with the 26 letters. This comparison is done such that it is very robust to the background cluster and deformation of the letters. In Figure 1.9 it matches each of the letter "j", "e", "l" and so on from the present set of 26 letters that are available.

4.2.2 Graph Construction
The second step after the matching of the letters from the 26 letters is the graph constructing step where the pairs of the letters are analysed to see whether they are consistent or can be used consecutively to form a word thus forming a graph.

4.2.3 Search Of Convincing Words In The Graph
From the graph that is constructed in the previous step the possible satisfying combination of real words is identified . However, most of them do not form real words as this only work for certain visual CAPTCHAs. The real words in the graph are selected and scores are assigned to them on the basis on how well their individual letters match the image [2].

4.3 Hacking Audio Based Captchas
In a recent survey the latest in a string of CAPTCHAs to be hacked by software is the Google's audio capture. The one cost effectiveness of hacking the audio captures that have not yet had automated systems developed is to pay less for a CAPTCHAs reading by humans. However the level of resources to achieve is required. More number of problems occur for the site operator

during the development of software to automatically interpret CAPTCHAs. The wider impact of this work might take some time to appear, but it provides an interesting proof of breaking audio CAPTCHASs [2].

4.4 Business In Captchas Hacking

Hacking the CAPTCHAs has become a popular illegal business in this era. The intelligent hackers hack the CAPTCHAs for a reasonable sum of money that could benefit both the hacker and their customer. They charge 2dollor for a correct solving of the CAPTCHAs. Usually these hackers publish their advertisement at the very own CAPTCHAs page to lure the customers if they want to get the access. By cyber security law the CAPTCHAs hacking is punishable crime which results in a loss of money for a site that is hacked [3].

5. FUTURE CAPTCHAS IDEAS

The recent type of CAPTCHAs is in use is the puzzle based CAPTCHAs where the pictures are divided into chunks and displayed in a random asking the user to identify the correct feasible image from that display. Although this is the latest it is used rarely in the sites. For the more secure and precise identification of human user the gaming CAPTCHAs comes in handy and CAPTCHAs can also be developed for the regional languages .This section shows some idea of regional language CAPTCHAs and future gaming CAPTCHAs that can be used in common that the bots can't play those games .

5.1 Captchas In Regional Language

As the CAPTCHAs is available only with the English and some numbers and special characters it can be proposed to develop in the regional languages so it can help the people who face difficulties with the common English language. For an example consider a regional language called devanagari script that is used in Sanskrit which is a old language in India can be used as a CAPTCHAs. In this case the letters of this script can be identified by the server by splitting it as shown in Figure 1.11 where it is classified as upper strip core strip and the lower strip that could be used in the identification of the letters [7].

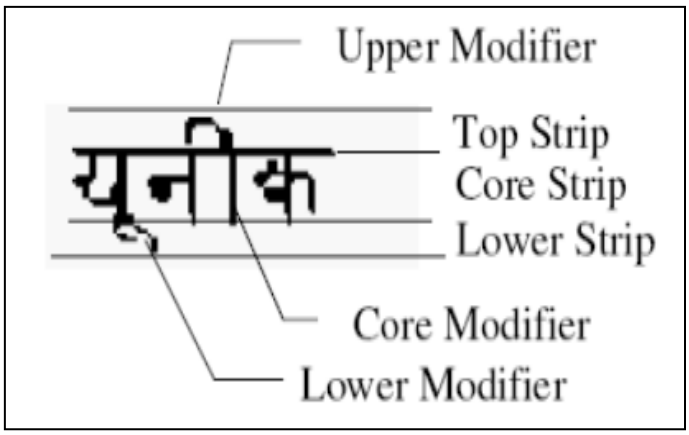

Figure 1.11 Devanagari script

The components of the proposed framework for the regional CAPTCHAs are given below [7].

5.1.1 Script's Database

The database consisting of sufficiently large data of devanagari script samples either in text or handwritten form .

5.1.2 Generating The Query

This phase is to develop a mechanism to query the database and obtain a random sample subject to the design

5.1.3 Obfuscation

This module takes the random sample from the database that distorts And adds noise to it.

5.1.4 Gui Of The Script

A user challenge-response interface

5.1.5 Match Respose

A determination of whether the user has submitted the accurate response for the challenge posed. Similarly these components and framework can be modified for the other regional languages such as Tamil,telugu,urudu,etc. But the process of striping the language remains the same. This can be complicated sometimes as there is more number of letters in the language unlike the English alphabets.

5.2 Game Based Captchas

As the demand for the strong CAPTCHAs is raising the Game based CAPTCHAs is a perfect solution to avoid the bots as the Game that is used here not only identifies the user but it sometimes entertains the user.

.

5.2.1 Dot Connector

Dot connector CAPTCHAs is a game based CAPTCHAs which is a type of a game that is easy and simple as the user is given a set of coloured dots and asked to connect either the same colour or the different .sometimes it is asked to connect the dots in a pattern that is mentioned. In Figure 1.12 the dot connector game is shown where the same red coloured dots are connected without touching the other colours [6].

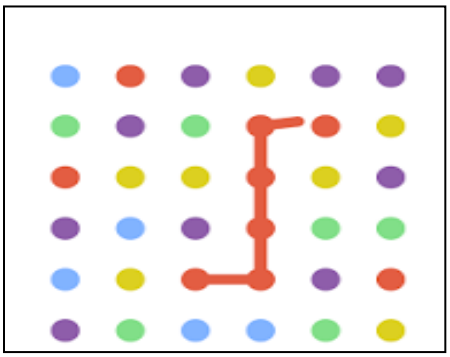

Figure 1.12 Dot connector CAPTCHAs

5.2.2 Shoot The Bird

Shooting the bird is exiting game that is given as a CAPTCHAs for the user to identify between the bots and the humans. This games asks to click the birds with the same colour to access the webpage this may be difficult for the bots. Figure 1.13 shows the shoot the bird game CAPTCHAs [6].

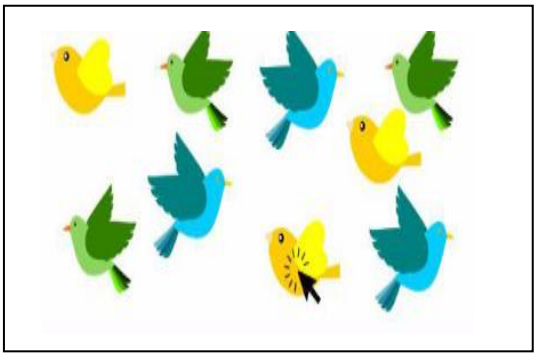

Figure 1.13 Shoot the bird CAPTCHAs

6. LIMITATIONS OF CAPTCHAS

Even though the CAPTCHAs is a secure barrier for the identification of the human user, it is often criticized for its limitations over the disable persons. For example say a person who is deaf and wants to access a webpage gets a audio based CAPTCHAs it makes difficult for that person to access that site. Similarly, if a person who accesses the site is illiterate he would find hard for identifying the CAPTCHAs, for example in mathematical CAPTCHAs if a person who is weak in solving problems cannot access the site.

7. CONCLUSION

In this paper, the different types of CAPTCHAs that are used in the sites are explained with the common attacks and some future ideas that can be implemented for the higher security. The usage of these CAPTCHAs has become very common in the website security starting from the letters and numbers to the games. As the CAPTCHAs types keeps on increasing the algorithms to crack these CAPTCHAs is also increasing. To ensure the security of these CAPTCHAs more advanced methods must be used like the game based or puzzle based CAPTCHAs which makes the access secure and also entertain the user. For more easy understanding the CAPTCHAs that belong to the respective regional languages will be more helpful for the user who do not know the English language. In future, more advanced and unbreakable types of CAPTCHAs may be developed for the necessity of the security [5].

8. REFERENCES

[1] Ved Prakash Singh, Preet Pal "Survey of Different Types of CAPTCHAS" (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2), 2014, 2242-2245, 2014.

[2] Greg Mori and Jitendra Malik. "Breaking a Visual CAPTCHAS.", Unpublished Manuscript, 2002.

[3] Sergei, Kruglov."Defeating of some weak CAPTCHASs" (http://www.captcha.ru/en/breakings/). Captcha.ru. . Retrieved 2008-12-21.

[4] Are You a Human. https://www.areyouahuman.com/.

[5] CAPTCHAS: http://www.CAPTCHAs.net/

[6] Mr. S. Ashok Kumar, Mr. N. Ram Kumar, Dr. S. Prakash, Mrs. K Sangeetha "Gamification of Internet Security by Next Generation CAPTCHASs" 2017 International Conference on Computer Communication and Informatics (ICCCI -2017), Jan. 05, 2017, Coimbatore, INDIA

[7] Sushma Yalamanchili and Kameswara Rao "A FRAMEWORK FOR DEVANAGARI SCRIPT-BASED CAPTCHAS "International Journal of Advanced Information Technology (IJAIT)., Vol. 1, No. 4, September 2011.

[8] Roman V. Yampolskiy and Venu Govindaraju. Embedded noninteractivecontinuous bot detection. Comput. Entertain., 5(4):7:1–7:11, March2008.

[9] Mohammad Moradi and Mohammad Reza Keyvanpour. CAPTCHA and its alternatives: A review. Security and Communication Networks,8(12):2135–2156, 2015.

[10] Luis von Ahn, Manuel Blum, and John Langford. Telling humans andcomputers apart automatically. Communications of the ACM, 47(2):56–60, February 2004.

[11] S. Saklikar and S. Saha. Public key-embedded graphic CAPTCHAs. In 2008 5th IEEE Consumer Communications and Networking Conference,pages 262–266, Jan 2008.

[12] Aditya Raj, Ashish Jain, Tushar Pahwa and Abhimanyu Jain "PictureCAPTCHAsWith Sequencing: Their Types and Analysis," InternationalJournal of Digital Society, vol. 1, no. 3, pp. 208-220, 2010.