HIGH SCHOOL EDITION
Journal of Student Research

# CAPTCHAs in reality: What might happen if they are easily broken?

Ananya Narayanaswamy[1] and Elaine Wah[#]

[1]Vidyashilp Academy, Jakkur, Bangalore, India
[#]Advisor

## ABSTRACT

This research paper explains the potential impact when CAPTCHAs become easy for attackers to solve. A literature review has been done on the different types of CAPTCHAs as well as the different methods present to break the CAPTCHAs. But data regarding the effect that CAPTCHAs have had when they have been attacked and broken is scarce. CAPTCHAs are deployed mainly to secure sites or apps but there are particular disadvantages which make them not only less beneficial to the owners of the sites or apps but also to the humans who need to solve them. These disadvantages are also in the form of vulnerabilities which prove beneficial to the attackers who can easily make use of the vulnerabilities to break the CAPTCHAs. It has been found that the impact of these attacks, in reality, has not been mentioned in previous studies. The impact prompts for advanced CAPTCHAs which solve the problems with the CAPTCHAs we have today.

## Introduction

CAPTCHA (Completely Automated Public Turing Test to tell Computers and Humans Apart) [1][2][3] is an automated Turing test also known as Human Interaction Proof(HIP)[6] in which computer programs ask the user to solve a particular challenge and grade the answer submitted by the user[4]. Its main purpose is to protect services and resources from automated misuse by malicious programs[5]. It checks whether the user is a human or robot by providing them with different challenges, like playing a game according to the instructions provided or selecting specific images. It is mainly used by websites or web-based email services in order to prevent spammers from creating several fake accounts using computer programs[6]. However, it has become easy for the attackers to find vulnerabilities in the CAPTCHAs which has led to many techniques and strategies used to attack the CAPTCHAs. Because CAPTCHAs are trusted to provide security to sites or apps containing users' data or have data that provides monetary value, CAPTCHAs are significant enough for the attackers to find issues in. This has led to attacks (elaborated more in section 3) which have had a significant effect both in a good (new techniques to develop CAPTCHAs are developed leading to new advances in the field) and a bad way (financial losses to the owner of the site or app where CAPTCHA is deployed).

This research paper gives an introduction to the CAPTCHAs, gives examples of the attacks that broke the CAPTCHAs and gives information about the potential impacts that CAPTCHAs can have when CAPTCHAs become easy for attackers to solve. The research paper is divided into five sections.Section 1 is the introduction, Section 2 holds information on the classification of CAPTCHAs, and different types of CAPTCHAs. Section 3 holds information on the Process of attacking text-based CAPTCHA, methods used by attackers to break other types of CAPTCHAs and examples of various attacks on different types of CAPTCHA. Section 4 holds information on the impact when CAPTCHAs become too easy for attackers to solve. Prior work has not studied the impact that less secure CAPTCHAs can have and papers reporting the losses incurred due to a CAPTCHA being broken have not been found. This research paper explains the impact in section 4.

## Classification of CAPTCHAs

This section gives an introduction to how CAPTCHAs have evolved, how they are classified and tells about the different schemes of CAPTCHAs which have been broken by a few attacks later in section 2. The idea of using CAPTCHA goes back to 1996[10]. Alta Vista developed a first practical example of a CAPTCHA scheme in 1997 to protect against fake account creations.In the year 2000, Von Ahn et al. introduced a few proposals to design CAPTCHA schemes which were easily solvable by humans but difficult for computer programs[11][2]. This year was also the year when the word 'CAPTCHA' was termed by John Langford, Nicholas J. Hooper and Luis Von Ahn at Carnegie Mellon University[12].

CAPTCHAs have been traditionally classified into six types.The different types are text-based, image-based, video-based, audio-based, math-based, and game-based CAPTCHA [7][8].However, a recent survey[9] has argued that CAPTCHA schemes today can be divided into ten categories: text-based, image-based, video-based, audio-based, game-based, math-based, slider-based, behaviour-based, and sensor-based CAPTCHAs. This new method of categorization was made as to the new CAPTCHA schemes such as reCAPTCHA V2 and Geetest do not fall under any category of the six traditional CAPTCHA schemes. This research paper will take both the above divisions into consideration in order to explain the types of CAPTCHAs. CAPTCHAs can be classified on a broad level into OCR and Non-OCR CAPTCHAs[21].OCR stands for optical character recognition which means converting the characters or the text written on the images into mechanical or electronic versions in order to be recognized by the computer. OCR CAPTCHAs are the optical character recognition CAPTCHAs which are the characters recognized by the computer software or devices. An example of these CAPTCHAs is the Text-Based CAPTCHAs which show a few CAPTCHAs with distorted letters or have some changes in their font colour/ size/ texture etc(to make them difficult for robots to identify). Non-OCR stands for Non-Optical Character Recognition CAPTCHAs. These types of CAPTCHAs are the CAPTCHAs where the characters are not shown for the user to recognize but other challenges are provided like solving a game, a puzzle or selecting the right image as asked by the CAPTCHA.

Different types of CAPTCHAs with a few examples of their schemes have been explained below.

### Text-based CAPTCHAS

Text-based CAPTCHA tests users by asking straightforward questions like" What is the addition of 2 and 3?" or asking the user to recognize the characters containing letters or digits or a mixture of both shown in the CAPTCHA. Text-based CAPTCHAs can be shown with the text in a distorted form as well as its original form. The distorted form of the text makes it more secure from being solved by the bots using OCR techniques. This is the most popular type of CAPTCHA, involving questions which are easily solvable by humans in a particular amount of time. They are easy to implement but require a huge database consisting of words. Sometimes the text is so small that it is difficult for the user to identify the characters in the CAPTCHA. These CAPTCHAs also require the user to be good at recognizing the text of a particular language which makes it harder for the users who are not acquainted with the language. There are different schemes of CAPTCHAS in text-based CAPTCHAS. Some are 2D text-based or 3D text-based whereas few others are animated text-based CAPTCHAs. A few examples of these three types of text-based CAPTCHAs are mentioned below.

Gimpy is an example of 2D text-based CAPTCHA which asks users to recognize three out of seven randomly selected words from the dictionary. Baffle text CAPTCHA is also a 2D text-based CAPTCHA that asks users to recognize a misrepresented word randomly selected from the dictionary. Figure 1a is an image of Baffle text CAPTCHA. The CAPTCHAs made by Microsoft (2002), Google (2006), Yahoo (2008), Megaupload.com (2010) [9], the handwritten CAPTCHA scheme proposed by Amelia Rusu and Venu Govindaraju[14], Ez-gimpy and clickable CAPTCHA are few more examples of 2D text-based CAPTCHAS. Teabag 3D, 3D CAPTCHA and Super CAPTCHA are examples of 3D text-based CAPTCHA which asks users to recognise a few 3D characters. Figure 1b is an image of SUper CAPTCHA. DotCHA is also an example of 3D text-based CAPTCHA which asks users to drag and rotate the image which contains the text to identify the characters in it.

Atlantis CAPTCHA is an animated text-based CAPTCHA which asks users to recognize characters which keep changing their font colour. Figure 1c is an image of Atlantis CAPTCHA. HelloCAPTCHA, Dracon CAPTCHA, Killbot professional and NuCAPTCHA are also a few examples of 3D text-based CAPTCHAS.

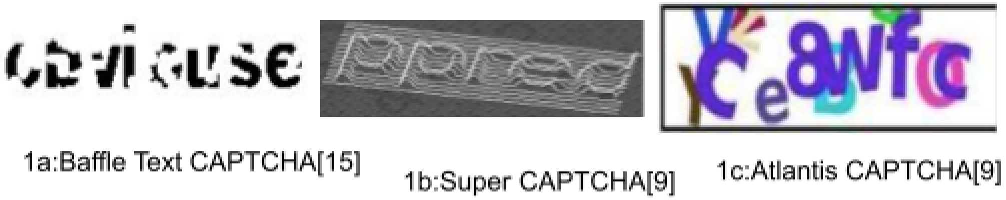
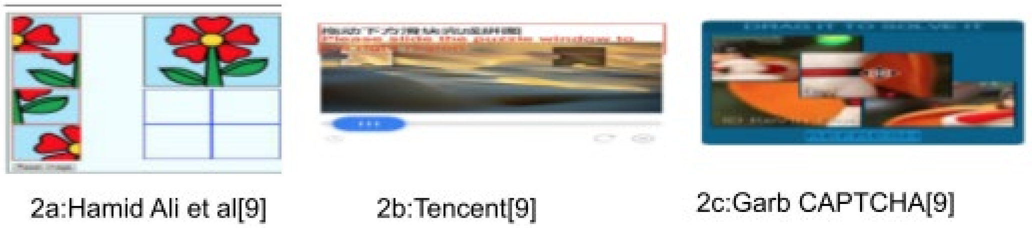

1a:Baffle Text CAPTCHA[15]
1b:Super CAPTCHA[9]
1c:Atlantis CAPTCHA[9]

**Figure 1.** 2D, 3D and animated text-based CAPTCHA

## Image-based CAPTCHAs

Image-based CAPTCHA is a CAPTCHA where the user is shown an image and is asked to do a particular task. There are different types of schemes in an image-based CAPTCHA. The CAPTCHA can ask the user to do a task depending on the type of scheme that the CAPTCHA belongs to. Though image-based CAPTCHAs are user friendly, their main issue is problems with image identification due to blurred images or low vision. Here are a few of the schemes that come under Image-based CAPTCHA.

Click-based CAPTCHA asks the user to click on a specific object or area in the image in order to complete the task. SACaptcha and Implicit CAPTCHA are examples of click-based CAPTCHA. Selection-based CAPTCHA asks the user to select particular images from a set of provided images. Facebook im-age CAPTCHA, HumanAuth, FR-CAPTCHAare, FaceDCAPTCHA, AVATAR and SEMAGE are a few examples of Selection-based CAPTCHAs. In the slid-ing image-based CAPTCHAs, the user is provided with a slider which needs to be used to solve the given challenge by adjusting, moving or correcting any part of the image as specified in the challenge. WHAT'S UP CAPTCHA, MintEye CAPTCHA and Tencent are a few examples of the Sliding image-based CAPTCHA. In Figure 2b we can see an image of Tencent. Drag and drop-based CAPTCHA requires the user to drag and drop parts of an image to fulfil a particular challenge given by the CAPTCHA (Few examples under this CAPTCHA scheme are considered puzzle-based CAPTCHA in other research papers). Garb CAPTCHA, Capy CAPTCHA, KeyCAPTCHA, CAPTCHA by Gao et al, and Hamid Ali et al are a few examples of drag and drop-based CAPTCHA. Figure 2c is an image of Garb CAPTCHA whereas Figure 2a is an image of CAPTCHA by Hamid Ali et al. Drawing-based CAPTCHAS require users to draw according to the challenge given. VAPTCHA, Drawing CAPTCHA and Motion CAPTCHA are a few Drawing-based CAPTCHAS. Interactive-based CAPTCHA similar to drawing based CAPTCHA relies on the interaction of the user. Users' movement can be using the mouse, hand or using his/her brain to solve a challenge. These movements cannot be made by robots, therefore, robots would not be able to solve the CAPTCHA. CAPTCHaStar and Cursor-CAPTCHA are a few Interactive-based CAPTCHAS.

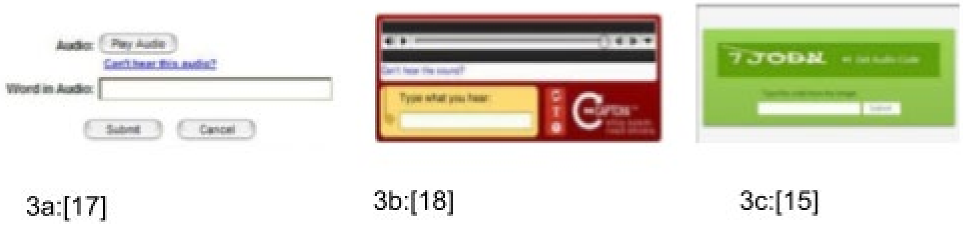


2a:Hamid Ali et al[9] 2b:Tencent[9] 2c:Garb CAPTCHA[9]

**Figure 2.** Examples of image-based CAPTCHA

## Audio-based CAPTCHA

Audio-based CAPTCHAs were introduced to act as a replacement for visual CAPTCHAS for visually impaired users. Users are required to listen to the audio clip and type the characters that they identify from the clip. Though audio-based CAPTCHAs are more difficult to break than the above CAPTCHAs, as audio-based CAPTCHAs use speech recognition, the audio might not be recognised by non-English speakers and the pronunciation might be similar for two or more characters. Here are a few schemes of Audio-based CAPTCHAs. Audio ReCAPTCHA asks the user to recognize the eight numbers which would be said in the audio clip. The clip would have background human noises. eBay audio CAPTCHA asks users to recognize six numbers which are spoken by people with different voices along with the same background noise for each of them. Audio reCAPTCHA(2013) asks users to identify all digits which are in the form of clusters with digits overlapping each other with each cluster containing three or four overlapping digits. Audio ReCAPTCHA(2017) asks users to recognize ten digits with background noise in the clip. Microsoft CAPTCHA is similar to e-bay CAPTCHA but has a background noise consisting of conversations. In Figure 3 we can see a few desktop images of audio-based CAPTCHA.

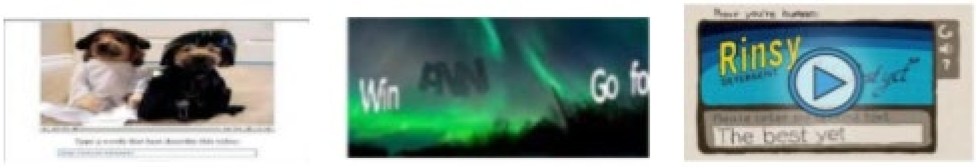


3a:[17] 3b:[18] 3c:[15]

**Figure 3**. Examples of desktop images for audio-based CAPTCHA.

## Video-based CAPTCHA

Video-based CAPTCHA provides users with a video. After watching the video, the user is asked a few questions based on the CAPTCHA scheme. The major limitation of this type of CAPTCHA is that the CAPTCHA is not practical because of the large size of the file but these CAPTCHAs are harder to break than the above-mentioned CAPTCHAs. Here are a few examples of Video-based CAPTCHA. Kluever el al asks a user to provide three random words that describe the video. Motion CAPTCHA asks the user to select a sentence that best describes the object in the video. The video can also ask the user to look at a few of the words which they would have to type down after the video finishes. In figure 4 we can see a few desktop images of video-based CAPTCHA.

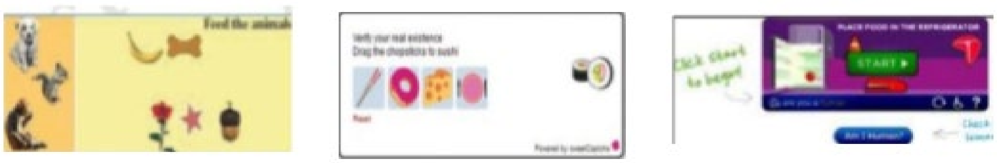


4a:[18] 4b:[19] 4c:[17]

**Figure 4.** Example of desktop images for video-based CAPTCHA

## Game-Based CAPTCHA

Game-based CAPTCHAs provide users with a game to be solved. It was created based on the assumption that humans are better at solving a game after understanding the challenges and the rules of the game. The limitation of this CAPTCHA is that the humans may not read the rules of the game or may not understand a game which is not provided with rules leading to the wrong steps in solving the CAPTCHA. Here are a few examples of Game-based CAPTCHA. A CAPTCHA called PlayThru is similar to the drag and drop CAPTCHAs where the user is required to move particular objects to a specific place. Sweet-CAPTCHA asks users to match specific images with the images provided in the game. Dice CAPTCHA asks the user to roll the dice with the help of a roll button after which the user is required to write the number that appears on the dice. Tic Tac Toe asks the user to play a game of Tic tac toe where the user is required to get a line having only three X's or three O's. DCG CAPTCHA and GISCHA are a few other types of Game-based CAPTCHA. In figure 5 we can see a few desktop images of game-based CAPTCHA.

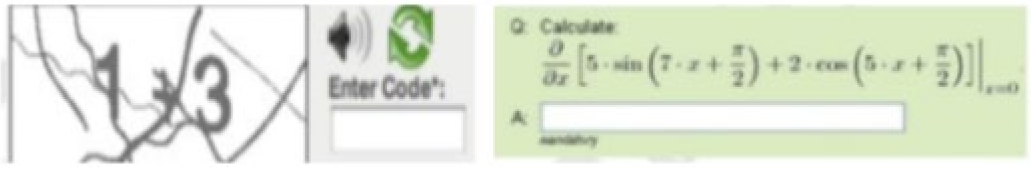


5a:DCG CAPTCHA[9] 5b:SweetCAPTCHA[9] 5c:PlayThru[17]

**Figure 5.** Example of desktop images for game-based CAPTCHA

## Math-based CAPTCHA

Math-based CAPTCHA requires the user to perform mathematical operations as asked in the challenge. The major problem here is that the user might not know how to perform mathematical operations or perform them incorrectly and therefore might not be able to solve the CAPTCHA. Here are a few examples of math-based CAPTCHA. Arithmetic CAPTCHA asks the user to enter the answer to a basic arithmetic operation which can include addition, subtraction, multiplication and division. QRBGS CAPTCHA asks the user to type the answer to a complex mathematical equation involving trigonometric and differential functions. In figure 6 we can see a few desktop images of math-based CAPTCHA.

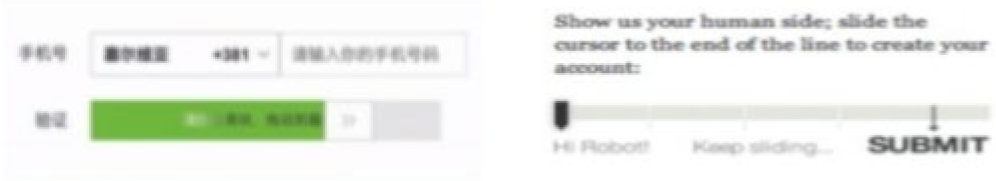

1x3
Enter Code*:
Q: Calculate:
$\frac{\partial}{\partial x}\left[5\cdot\sin\left(7\cdot x+\frac{\pi}{2}\right)+2\cdot\cos\left(5\cdot x+\frac{\pi}{2}\right)\right]\Big|_{x=0}$
A:
mandatory

6a:Arithmetic CAPTCHA[9] 6b:QRBGS CAPTCHA[9]

**Figure 6.** Example of desktop images for math-based CAPTCHA

## Slider-based CAPTCHA

Slider-based CAPTCHA requires the user to move the slider in order to prove that they are human. It is mainly adopted because it is easy for humans to solve irrespective of how easy or difficult it is for computer programs to solve. Here are a few examples of slider-based CAPTCHA. Taobao asks users to drag a slider from its starting point to its ending point. TheyMakeApps CAPTCHA asks the user to drag the slider to its ending point just like the above-mentioned CAPTCHA. In figure 7 we can see a few desktop images of video-based CAPTCHA.



Show us your human side; slide the cursor to the end of the line to create your account:

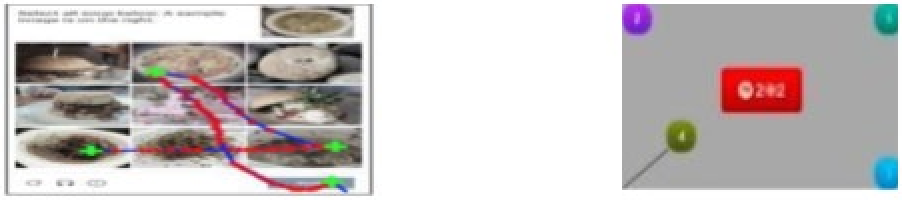
7a:Taobao[9]

7b:TheyMakeApps[20]

**Figure 7.** Example of desktop images for slider-based CAPTCHA

## Behaviour Based CAPTCHA

Behaviour-based CAPTCHAS recognize humans by their physical characteristics by taking into consideration their swipe, mouse, keystroke dynamics and eye movement. A few examples of behaviour-based CAPTCHAS include BeCAPTCHA-Mouse which asks the user to solve CAPTCHA based on an image which is similar to the question asked by reCAPTCHA V2 but in BeCAPTCHA-Mouse the user's mouse trajectory is also taken into consideration when the user is solving the CAPTCHA. Be-CAPTCHA is another example of Behaviour-based CAPTCHA where the user is required to solve a slider based challenge but here too the users' sensor data and swiping gestures are taken into consideration. EYE-CAPTCHA, Invisible reCAPTCHA, GEEtest, and NETease are a few more examples of behaviour-based CAPTCHA. A few examples of desktop images of behaviour based CAPTCHAs are shown in Figure 8 below.



8a:BeCAPTCHA-Mouse[9]

8b:Eye-CAPTCHA[9]

**Figure 8.** Example of desktop images for behavior-based CAPTCHA

## Sensor-Based CAPTCHA

Sensor Based CAPTCHA uses hardware sensors to recognize whether a user is a human or a bot. There are physical sensor-based CAPTCHAS where the user is required to do a physical task like taking a short walk of a few steps or tilting the mobile phone. For example, Sensor CAPTCHA asks the user to move the device and Pedometric CAPTCHA asks the user to walk at least five steps. There are also cognitive sensor-based CAPTCHAS where the user solves a cognitive challenge just like how the user does in the traditional challenges (playing a game, selecting images etc) but here the user needs to use the sensors present in the CAPTCHA instead of the usual swiping or taping. For example, SenCAPTCHA asks the user to move a red ball that they would provide into the centre of a picture of an animal's eye by tilting the device. A few examples of desktop images of audio-based CAPTCHAs are shown in Figure 9 below. [9]

Figure 10 provides a schematic visualization of the different types of CAPTCHAs.

First you need to walk 5 steps in order to sign in.

Steps:
5
Congratulations, you have walked 5 steps.
Username: Walk

9a:SenCAPTCHA[9]

9b:Pedometric CAPTCHA[9]

Figure 9: Examples of desktop images for
Sensor-based CAPTCHA

**Figure 9.** Example of desktop images for sensor-based CAPTCHA

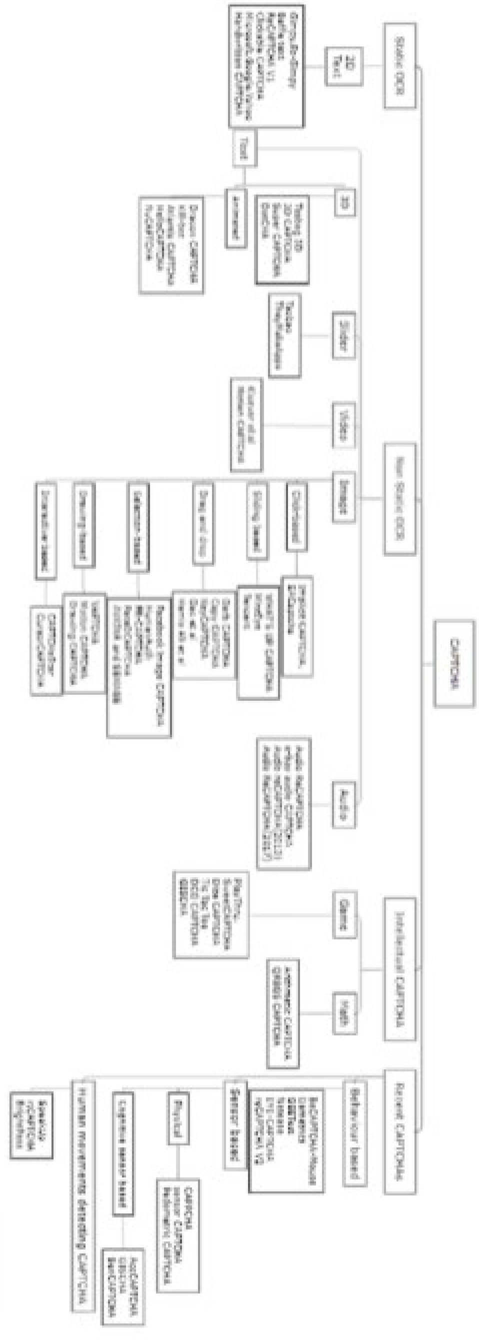


Figure 10: Types of CAPTCHA

**Figure 10.** Types of CAPTCHA

## Methods of attacks on different types of CAPTCHAs

This section explains the common methods used to attack the text-based CAPTCHA schemes (which are widely used schemes compared to the other types of CAPTCHAs), and explains the methods using which the attackers try to find vulnerabilities in the CAPTCHA (in order to attack different types of CAPTCHAs) and gives examples of the attacks that make use of different vulnerabilities in the CAPTCHAs to break them. The information in this section is written to show how attackers have already made use of the insecure CAPTCHA systems and how they have attacked the CAPTCHA schemes successfully. This raises a question about the impacts when the attacks are used on the CAPTCHAs in the real world today which is explained further in the paper.

Attackers have been making use of various methods of attack to break CAPTCHAs and the methods have been evolving to such an extent that a proposal made by Fabian Stark, Caner Hazırbas, Rudolph Triebel, and Daniel Cremers uses active learning to train the computer program to recognize CAPTCHAs and it does not require human interference[22]. The computer program learns to solve the CAPTCHAS by itself by learning from its mistakes and storing the steps it took when it solved the CAPTCHAs correctly.

There are many techniques and strategies used in an attack and sometimes a single computer program can break more than one type of CAPTCHA whereas the other times it can only break a specific type of CAPTCHA. Most of the time a new CAPTCHA is developed after finding the vulnerabilities of the previous CAPTCHAs. Attackers try to find vulnerabilities in the newly created CAPTCHA and break them which forces the CAPTCHA researchers to address the vulnerability and create CAPTCHAs which are more secure and usable.

According to the authors in [9] the methods used by the attackers fall under these categories: Object Recognition attacks, Random guess attacks and Hu-man Solver Relay attacks. Object Recognition attacks normally include pixel count (number of pixels in an image), database attacks (an attack on the database which allows complete authentication by-pass to the attacker allowing him/her to extract any data from the database) and dictionary attacks (attackers try to guess passwords to break into accounts) and the techniques used are OCR recognition, pattern matching (according to DeepAI[12], it is an algorithmic task that finds predetermined patterns among sequences of raw data or processed tokens), Scale-Invariant Feature Transform (an algorithm used to detect local features in images), and Deep learning (based on neural networks and uses multiple layers of processing to extract high-level features from data). Random guess attacks consist of the attacker trying to guess the correct answer to break the CAPTCHA. Human Solver Relay attacks consist of humans in remote areas who solve the CAPTCHAS for a very small income. Few methods of attack are explained below.

A bypass attack is an attack when the attacker does not need to solve the CAPTCHA and the solution is made visible to the attacker accidentally. This attack does not belong to any of the three methods of attacks explained above, since object recognition attacks techniques, guessing or using human solvers have not been used to solve this attack. A Challenge Replay attack is an attack where the system has a small database of CAPTCHA challenges which can be taken note of and answered by a human. Further, those solutions to the CAPTCHAS are saved by the automated agent to solve the CAPTCHAS. Signal Processing involves CAPTCHA researchers and attackers making reversing CAPTCHA trans-formations via mathematical heuristics (strategies used to solve complex word problems) and machine learning approaches. For Text-Based CAPTCHAS, the process of processing, segmentation and recognition mentioned above is used. Mechanical Turk attack is similar to a challenge replay attack but here as soon as the automated agent receives a CAPTCHA to which there is no solution in its database, the automated agent sends it to the human agent who is paid to solve the CAPTCHAS full time just like how humans are taken for labour in the human solver relay attacks. The human agent solves the CAPTCHA and types the correct answer along with saving the solution in its database. A trivial Guessing attack is used when there is an unlimited range of challenges and a very limited number of answers to a particular question (for example: which of these 20 options is correct). An attacking program is programmed in such a way that it randomly guesses all the options of the answers until it gets a correct solution to the challenge. This attack can only be used when the CAPTCHA has a wide error tolerance for user inaccuracy. A Brute Force attack is an attack which is used when there is limited space for inputting the answer (For example, a CAPTCHA asking for a 3 digit answer). This attack uses a group of automated agents which guess an answer (For example: if an answer requires three digits as its input, the automated agents try all the digits from 100 to 999 until the CAPTCHA verifies the agent as human). A hybrid attack is a combination two or more of the above attacks to solve a particular type of CAPTCHA.[25]

Since most of the CAPTCHAS used are Text-Based CAPTCHAS attackers normally follow the methodology of preprocessing, segmentation and recognition to break text-based CAPTCHAS. Preprocessing consists of steps which make it easier to extract the characters in the next step(segmentation). Preprocessing includes removing the background, upsampling (dividing a pixel into sub-pixels), blurring, thresholding (removing pixels of low intensity), and line removal (removing lines which are not a part of characters), thinning of characters and identifying the CAPTCHA scheme. The next step is called segmentation. It locates character from the image and separates it from the image. The last step is called character recognition. Classifiers are used to identify the characters. Machine learning algorithms can be trained to recognize the output obtained from the second step[24].

Further this section gives examples of attacks on different types of CAPTCHAs. Almost all the CAPTCHAs have been broken by the attackers as shown in Figure 11. This implies that the CAPTCHAs have to be continuously improvised before the attackers can find a way to break them. A few examples of attacks show that one attack (For example, DeCAPTCHA ) can easily break one or more types of CAPTCHAs which shows how easy the CAPTCHAs have become for robots to solve. It is mentioned below about CAPTCHAs being broken by attack-ers which raises a question about the security of using insecure CAPTCHAs and how this might affect the usage of CAPTCHAs in the future. Since the CAPTCHAs are too easy for the attacker to solve, the next section would address the potential impacts of attackers being able to break CAPTCHAs easily using a few real-world examples.

## Attacks against Text-Based CAPTCHA

Mori and Malik proposed an attack which could break both Gimpy and Ez-Gimpy with 92 per cent and 33 per cent accuracy[26]. Later Ez Gimpy was also broken with a 99 percent accuracy [9]. In 2011, Vu Duc Nguyen, Yang-Wai Chow and Willy Susilo proposed a way to break Teabag 3D which is a 3D CAPTCHA Text-based CAPTCHA[27]. Thirteen out of Fifteen Text-Based CAPTCHA schemes from top websites were broken by a single bot in an experiment conducted by Elie Bursztein, Matthieu Martin, and John C. Mitchell[24] which gives an idea of how dangerous a strong real-world attack on the CAPTCHAs can be. This experiment was conducted on Authorize, Baidu, Blizzard, Captcha.net, CNN, Digg, eBay, Google, Megaupload, NIH, Recaptcha, Reddit, Skyrock, Slashdot, and Wikipedia out of which Google and Re-captcha resisted their attempts. A case study showed an attack that could break text-based CAPTCHAS with nearly 100 per cent accuracy using a pixel-count attack[29] which is another example for the user to be cautious about the CAPTCHA that is deployed on his/her site, app, etc.

## Attacks against Image-Based CAPTCHA

A side-channel attack proposed by Hernandez-Castro et al. broke the Hu-manAuth CAPTCHA challenge with a 92 percent success rate[30]. Google and Facebook image CAPTCHAS were broken with a success rate of 70.78 per cent and 83.5 percent. The attacks created by authors in [31] could break different schemes of Image-based CAPTCHA at a 100 percent success rate. This survey [32] used German Software Development Analytics CAPTCHA Breaker and Tesseract to break 337 Image-Based CAPTCHAS. The drawing CAPTCHA was broken by Lin et al. with an accuracy of 75 per cent [9].

## Attacks against Audio Based CAPTCHA

Though audio-based CAPTCHAs were thought to be difficult to break for computer programs, the evolution in technology has made many successful attacks on them possible. A survey used three attacks: Tesseract, GSA CAPTCHA Breaker, and DeCAPTCHA to break four audio-based CAPTCHAS. In[33] the authors proposed a CAPTCHA solver that broke the Microsoft and Yahoo audio CAPTCHAS. DeCAPTCHA was also used to break nearly 75 per cent of the eBay audio based CAPTCHA[34]. Besides just breaking the CAPTCHAs, research shows a new way of using which attackers can easily at-tack on a low cost and the cost of making them reduced on using active and supervised learning instead of the expensive automated attacks as proposed in [35].

## Attacks against Video, Math, Game, Slider Based CAPTCHA

Not only the above CAPTCHAs but nearly all the CAPTCHA schemes have been broken by the attackers. Here are a few examples of attacks on other CAPTCHAS. An example of breaking Video-Based CAPTCHA can be seen in [36] where the author uses Tesseract to break the video-based CAPTCHA. Using image processing techniques and unsupervised learning, this study [37] was able to break game-based CAPTCHA. Different variations of slider based CAPTCHA were broken using a simple JavaScript code and a puppeteer in [38]. QRBGS CAPTCHA was broken using a side-channel attack by Hernandez-Castro et al. with a success rate of 44.45 per cent in [39]. SweetCAPTCHA was broken by a solver in [40].

## Attacks against behaviour-based CAPTCHA

An attack proposed against GEEtest and Netease CAPTCHA schemes broke the schemes successfully with success rates of 96 per cent and 98 per cent by using the Sigmoid function in [9]. In [41] the authors broke reCAPTCHA V2 and further proposed how to break the second layer of reCAPTCHA V2 in with a success rate of 70.78 per cent. In [42] authors defeated the No CAPTCHA reCAPTCHA scheme for any grid resolution.

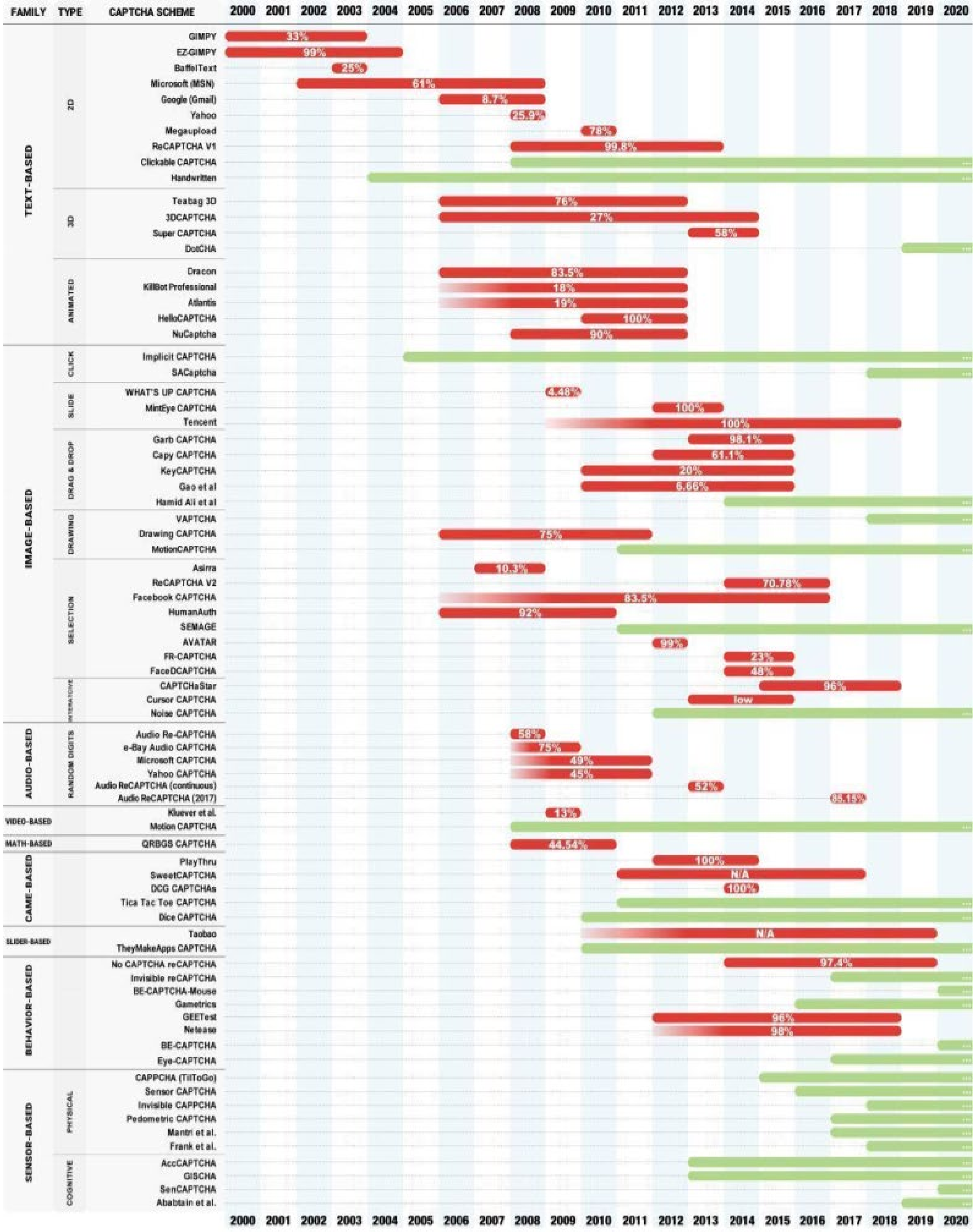


**Figure 11.** Attacks found for all the CAPTCHAS and best breaking percentage received by them [9]

## Impact when CAPTCHAs become easy for attackers to solve

As said in the prior sections, nearly all the types of CAPTCHAs that have been implemented have been broken[9]. There are methods using which most of the CAPTCHA schemes can be broken. With improvement in CAPTCHAs, there has been the development of strategies and techniques using AI, Deep learning and more to break the CAPTCHAs. It can seem like CAPTCHAs are not a very significant part of a site, app or any other means where it is deployed but when a CAPTCHA which is less secure is deployed on a site, it can have very large impacts like spam messages[49], credential stuffing[50] and misuse of resources[47] which can cause losses to the person owning the site, app, etc.

In this section, the paper explains the benefits and losses that people have experienced due to less secure CAPTCHAs being attacked and broken successfully on their sites and apps (like the examples of attacks experimented by the researchers in section three). The section would also explain the potential impact when CAPTCHAs become too easy for attackers to solve along with a few real world examples of successful attacks on insecure CAPTCHAs in the past.

Unfortunately, prior works have not studied the specific schemes of CAPTCHAs that were broken in the real world by the attackers. Because of this, only attitudinal, economic, and training impacts have been explained below using real world examples of impact when attackers had broken a few CAPTCHAs.

### Attitudinal impact

Attitudinal impact is an impact which occurs when there is a change in attitude. Here there is a change in the attitude of CAPTCHA users who feel that it might be better to switch to other authentication systems instead of using CAPTCHAs. They feel that different methods that have a stronger security system and frustrate humans less should be deployed instead of the less secure CAPTCHAs that have been used till now.

A research paper[43] reported practical attacks on three e-banking CAPTCHA schemes for transactional verification and attacks on forty-one schemes for lo-gin. These CAPTCHAs were deployed by many countries in the world including Germany, the USA, China, and countries in Asia, Europe, North America etc. The authors found that all the 41 CAPTCHA schemes used for login were insecure against automated segmentation attacks. They are of the opinion that CAPTCHAs might be incapable of protecting the e-banking system as a single attack can potentially cause huge losses to both the members of the banks and the banks themselves. They further suggest moving to alternative solutions like hardware security tokens which they feel are more promising despite their disadvantages (some hardware security tokens cannot resist man-in-the-middle attacks where the attacker listens to the communication between the user and the host). Articles like [44] by Kasada and [45] by Perimeter have stated that humans are becoming frustrated with CAPTCHAs (as they think it is time-consuming and discourages them to use the site where CAPTCHA has been deployed) and that using CAPTCHAs to check a human if they are a human needs to stop. sites like SAASPASS [46] have introduced CAPTCHAs to be replaced entirely with “robust two-factor authentication“ which uses a bar-code scanning system, One time passwords, and remote login (where users can log in with just a click on their mobile phones) etc. This is less time consuming for humans and secure as bar-codes, passwords etc are different for each user, therefore, reducing mass attacks on a large number of devices.

### Economic impact

Economic impact happens when there are monetary benefits or losses because of the situation. Because of attacks on CAPTCHAs, there have been economic losses for the site and companies using CAPTCHA. Here are a few examples.

Alibaba, a Chinese company suffered an economic loss of nearly 3 billion during the 2017 Double 11 Shopping Carnival because of the underground CAPTCHA solving services through which miscreant users obtained coupons and bargain-priced goods with one individual gaining an economic benefit of 32000 dollars[47]. After doing an economic analysis on a few CAPTCHA solving services (CAPTCHA solving services employ humans to solve the CAPTCHAs for a specific price per each CAPTCHA they solve) a research study [47] found that the daily revenue of ruokuai, a CAPTCHA solving service is about 3.6 million. The daily income of yundama is about 1.85 million, hyocr is about 0.29 million, dama2 is about 1.62 million, and AntiCAPTCHA about 0.10 million. The research paper also stated that the whole underground CAPTCHA solving market can be estimated at over 162.64 million, given that there are at least 152 active CAPTCHA solving services. The authors of [48] have said that the CAPTCHA solving services were mainly focused on the countries with low labour costs and less cost for internet services. These countries included India, China, Vietnam etc. So, if CAPTCHAs become too easy for robots to solve, it would lead to less CAPTCHA solving services involving human labour and more computer vision and image recognition services to solve the CAPTCHAs. This would make many people (potentially more than 450,000 [47]) unemployed. Using spam messages, companies or sellers advertise their products and services to the user and sometimes when the user uses their services or products and visits their websites, the user is doing a favour to these advertisers by being influenced to buy their products or services. Research stated that there was an increase in spam messages as soon as the LiveMail accounts' CAPTCHA system was reported to be broken[49]. A report stated that an attack on Imperva lasted 60 hours and included 44 million login attempts which was the cause of credential stuffing attacks targeted on a single company[50].

## Impact on CAPTCHA development

There is an impact on CAPTCHA development when new CAPTCHAs have to be created which are better in terms of security and usability than the CAPTCHAs we have today. As CAPTCHAs are becoming too easy for the computer programs to attack, there have been several proposals and ideas which can be used instead of the CAPTCHA schemes used today. Here are a few examples.

A research paper[51] proposes a new way to detect computer programs by using CAPTCHAs which asks the user to recognize a voice among other human voices ( a phenomenon called the Cocktail Party Problem ) which is easier for humans to recognize than the computer programs. Another research paper introduces a design of an audio CAPTCHA system which can be recognized with high probability by computer programs but can be recognized with low probability by humans[52]. This design tests if the user is a robot instead of checking if the user is a human. A new CAPTCHA scheme called GAPTURE was proposed to improve the strength of image-based CAPTCHAs against automated solvers by the authors ( Dorjan Hitaj, Briland Hitaj, Sushil Jajodia, and Luigi V. Mancini ) of [53]. They also conducted experiments to check how hard their CAPTCHA is for robots to solve. A new CAPTCHA called iCAPTCHA was proposed by the authors ( Huy D. Truong, Christopher F. Turner, Cliff C. Zou) of [28]. They show the vulnerabilities of existing CAPTCHAs by developing IMCA ( Instant Messanger CAPTCHA Attack ) and propose iCAPTCHA along with performing 3rd party human CAPTCHA attacks on iCAPTCHA.

As seen in section three, almost all the CAPTCHAs were broken long ago and even now researchers are breaking the CAPTCHAs with new attacks that are quicker and more accurate. If this continues and new CAPTCHA schemes do not get implemented, the attackers are sure to make use of this vulnerability. With more misuse of CAPTCHAs would come monetary losses on a large scale to the people who have deployed CAPTCHAs as their main security.

Though the examples of the impacts that CAPTCHAs which are easy can have is very less, the losses suffered by the targets of the attackers [47] [50] [49] are too large to ignore. It is necessary for people to strengthen the CAPTCHAs that they have been using or use an entirely different authentication system or try a combination of two or more schemes of CAPTCHA to make it difficult for the attackers to solve.

## Conclusion

CAPTCHAs have been used for authentication purposes for a very long time as said in section two and might not be used as widely as it is used today because of new authentication methods which are far more secure, usable and practical than CAPTCHAs. The CAPTCHAs must either be made more secure, harder for the robots to solve and easier for humans to solve or other authentication methods would have to be used instead of the less secure CAPTCHAs which frustrate the humans.

CAPTCHAs which are automated Turing tests to tell humans and computer programs apart are of various types. It includes text-based CAPTCHA, Im-age based CAPTCHA, audio-based CAPTCHA, video-based CAPTCHA, game-based CAPTCHA, math-based CAPTCHA, slider-based CAPTCHA, behaviour based CAPTCHA and sensor-based CAPTCHA. Usually, the process of preprocessing, segmentation and recognition (explained in section 3) are followed to break the text-based CAPTCHAs. Various methods can be used by the attackers to break the CAPTCHAs which include by-pass attacks, challenge replay at-tacks, mechanical Turk attacks, brute force attacks, human solver relay attacks, signal processing, hybrid attacks etc. The techniques to break the CAPTCHAs have evolved simultaneously with the techniques and strategies used to break the CAPTCHAs. CAPTCHAs are used mainly for security against attackers misusing the site or app resources, gaining access to users' credential details or trying to advertise a service or site in exchange for financial gains. When at-tacked successfully the are losses and gains which have been explained in terms of three major impact that were found from previous studies and reports when CAPTCHAs were attacked in reality. The three major impact (explained in section 4) are the attitudinal impact, economic impact, and CAPTCHA development impact. The attitudinal impact is the change in CAPTCHA users' attitudes toward using other authentication systems instead of CAPTCHAs as CAPTCHAs have become less secure and more time-consuming. The economic impact is an impact which occurs where there are financial gains or losses. CAPTCHA development impact is an impact which happens when there are new advances in techniques and strategies used to create new CAPTCHAs which are far more stable, difficult for humans to solve and easier for humans to solve.

## Acknowledgments

I would like to thank my mentor, Elaine Wah who provided me with guidance and support throughout the research. I would also like to thank my family for supporting me through long nights of research. This would not have been possible without you!

## References

Ahn, L. von, Blum, M., and Langford, J. 2003. Telling humans and computers apart automatically. Comm. of the ACM. 46 (Aug. 2003), 57-60
http://www.cs.cmu.edu/afs/cs/Web/People/aladdin/papers/pdfs/y2004/captcha_cacm.pdf

Ahn, L. von, Blum, M., Hopper, N. J., and Langford, J. 2003. CAPTCHA: Using hard AI problems for security. Eurocrypt'2003 https://link.springer.com/content/pdf/10.1007%252F3-540-39200-9_18.pdf

Baird, H. S. and Popat, K. 2002. Human interactive proofs and document image analysis. In Proc. of Document Analysis https://link.springer.com/content/pdf/10.1007/3-540-45869-7_54.pdf

Attacks and Design of Image Recognition CAPTCHAs Bin B. Zhu*, Jeff Yan, Qiujie Li, Chao Yang, Jia Liu, Ning Xu, Meng Yi, Kaiwei Cai Systems 2002. 507–518. https://www.researchgate.net/profile/Bin-Zhu/publication/221609266_Attacks_and_design_of_image_recognition_CAPTCHAs/links/00463517f8f8921e7d000000/Attacks-and-design-of-image-recognition-CAPTCHAs.pdf

CAPTCHA Smuggling:Hijacking Web Browsing Sessions to Create CAPTCHA Farms Manuel Egele, Leyla Bilge, Engin Kirda, Christopher Kruegel http://seclab.nu/static/publications/sac2010captcha.pdf

Are you a Human or Robot? or Everything CAPTCHA Revati Ghadge, Archana M. Naware http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.736.374&rep=rep1&type=pdf

Darko Brodi´c and Alessia Amelio. 2020. Types of CAPTCHA. Springer International Publishing, Cham, 29–32.

Ved Prakash Singh and Preet Pal. 2014. Survey of different types of CAPTCHA. International Journal of Computer Science and Information Technologies 5, 2 (2014), 2242–2245. https://d1wqtxts1xzle7.cloudfront.net/33706358/ijcsit20140502289-with-cover-page-v2.pdf?Expires=1652591625&Signature=FACK6Ukqh2OJ2BpnknWu4I6Qi4N1ywBhpFVKlJgsouNp5ATHKwWqg2lSB4tcreDmVQHT4xRcx0G-qa6O4kQpzvZ5PW3DPwYRqbKa5OMUi-Qnoih-vwBQGrLFfEelovtO8tJFfl7pVsFgQuTFg7eT5Z7iyhzDhuXuo0k0uYY1qYew--4JKhxktRbg7-pmwAq7lcIO1D0US3MbX62vjnooeJS3ZFfeDdnKuUEtMnZoLNZGv~bpMB7~9~z1jZC2i0Zd91aYCDlDf~iQmy4-05LUG-mD2BIih0aEK6cJL-JrO9L1kkO3YMO95~Mh4ummUk5RqHvouBENMe8Z4R53HWjfbA__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

Gotta CAPTCHA ’Em All: A Survey of Twenty years of the Human-or-Computer Dilemma. https://arxiv.org/pdf/2103.01748.pdf

.Moni Naor. 1996. Verification of a human in the loop or Identification via the Turing Test. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.50.6383&rep=rep1&type=pdf

Luis von Ahn, Manuel Blum, Nicholas Hopper, and John Langford. 2000. CAPTCHA: Telling Humans and Computers Apart Automatically. http : //www.captcha.net/

https://deepai.org/machine-learning-glossary-and-terms/pattern-matching

Ahn L. von, M. Blum and J. Langford. 2004. Telling Humans and Computer Apart Automatically. Communications of the ACM. 47(2): 57-60.

A. Rusu and V. Govindaraju. 2004. Handwritten CAPTCHA: using the difference in the abilities of humans and machines in reading handwritten words. In Ninth International Workshop on Frontiers in Handwriting Recognition. 226–231. https://doi.org/10.1109/IWFHR.2004.54

A REVIEW ON CAPTCHA GENERATION AND EVALUATION TECHNIQUES Mir Aman Sheheryar, Pradeep Kumar Mishra and Ashok Kumar Sahoo http://www.arpnjournals.org/jeas/research_papers/rp_2016/jeas_0516_4187.pdf

A comprehensive Study for Different Types of CAPTCHA Methods and Various Attacks Menna M.Elbalky, Medhat A. Tawfeek and Hamdy M. Mousa https://www.researchgate.net/profile/Menna-Magdy-11/publication/353225639_Issue_6_wwwjetirorg_ISSN-2349-5162/links/60ee1ed39541032c6d39edf3/Issue-6-wwwjetirorg-ISSN-2349-5162.pdf

DIFFERENT TYPES OF CAPTCHA: A LITERATURE SURVEY Vishal Shinde1, Prof. Vijay Rathi2 http://www.oaijse.com/VolumeArticles/FullTextPDF/185_11.DIFFERENT_TYPES_OF_CAPTCHA__A_LITERATURE_SURVEY.pdf

AN EVALUATION OF DIFFERENT TYPES OF CAPTCHA: EFFECTIVENESS, USERFRIENDLINESS, AND LIMITATIONS Karmand H. Abdalla, Mehmat Kaya https://www.researchgate.net/profile/Karmand-Hussein/publication/340006275_AN_EVALUATION_OF_DIFFERENT_TYPES_OF_CAPTCHA_EFFECTIVENESS_USER-_FRIENDLINESS_AND_LIMITATIONS/links/5e728823a6fdcc37caf62ccf/AN-EVALUATION-OF-DIFFERENT-TYPES-OF-CAPTCHA-EFFECTIVENESS-USER-FRIENDLINESS-AND-LIMITATIONS.pdf

A SURVEY ON THE DIFFERENT IMPLEMENTED CAPTCHAS Shadi Khawandi, Firas Abdallah and Anis Ismail https://www.researchgate.net/profile/Firas-Abdallah/publication/330827684_A_Survey_On_The_Different_Implemented_Captchas/links/5d85f24c458515cbd1a572ee/A-Survey-On-The-Different-Implemented-Captchas.pdf

Luke Wroblewski. 2010. A Sliding Alternative to CAPTCHA? https://www.lukew.com/ff/entry.asp?1138

Question-Based CAPTCHA by Mohammad Shirali-Shahreza and Sajad Shirali-Shahreza https://www.researchgate.net/profile/Mohammad-Shirali-Shahreza/publication/4309155_Question-based_CAPTCHA/links/53d0ef0f0cf25dc05cfe73be/Question-based-CAPTCHA.pdf

CAPTCHA Recognition with Active Deep Learning Fabian Stark, Caner Hazırba¸s, Rudolph Triebel, and Daniel Cremers http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.710.1085&rep=rep1&type=pdf#page=94

Character Segmentation for Automatic CAPTCHA Solving Christos Makris and Christopher Town https://benthamopen.com/contents/pdf/COMPSCI/COMPSCI-1-1.pdf

E. Bursztein, M. Martin, and J.C. Mitchell, “Text-based CAPTCHA strengths and weaknesses”, In: Computer and Communications Security(CCS), October 2011. http://www.decom.ufop.br/menotti/rp142/sem/sem1-dp3-artigo.pdf

Strengthening CAPTCHA based websecurity by Graeme Baxter Bell. https://researchrepository.murdoch.edu.au/id/eprint/8064/1/CAPTCHA-based_Web_security.pdf

G. Mori and J. Malik. 2003. Recognizing objects in adversarial clutter: breaking a visual CAPTCHA. In 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., Vol. 1. I–I.
http://wang.ist.psu.edu/imagination/mori.pdf

Breaking a 3D-based CAPTCHA Scheme Vu Duc Nguyen, Yang-Wai Chow and Willy Susilo https://www.researchgate.net/profile/Yang-Wai-Chow/publication/262372314_Breaking_a_3d-based_CAPTCHA_scheme/links/5f7fc336a6fdccfd7b51d4ea/Breaking-a-3d-based-CAPTCHA-scheme.pdf

iCAPTCHA: The Next Generation of CAPTCHA Designed to Defend Against 3rd Party Human Attacks http://cs.ucf.edu/~czou/research/iCaptcha-ICC2011.pdf

CAPTCHA Security by Jeff Yan and Ahmad Salah El Ahmad
https://eprints.ncl.ac.uk/file_store/production/152438/3F53CA83-C049-4D60-85F2-F66AAB7057EB.pdf

C. J. Hernandez-Castro, A. Ribagorda, and Y. Saez. 2010. Side-channel attack on the HumanAuth CAPTCHA. In 2010 International Conference on Security and Cryptography (SECRYPT). 1–7.
https://www.scitepress.org/papers/2010/29940/29940.pdf

Binbin Zhao, Haiqin Weng, Shouling Ji, Jianhai Chen, Ting Wang, Qin-ming He, and Reheem Beyah. 2018. Towards Evaluating the Security of Real-World Deployed Image CAPTCHAs. In Proceedings of the 11th ACM Workshop on Artificial Intelligence and Security (Toronto, Canada) (AISec '18). Association for Computing Machinery, New York, NY, USA, 85–96. https://doi.org/10.1145/3270101.3270104

LARGE SCALE CAPTCHA SURVEY by Mecheal Greene
https://udspace.udel.edu/bitstream/handle/19716/23980/Greene_udel_0060M_13475.pdf?sequence=2&isAllowed=y

E. Bursztein, R. Beauxis, H. Paskov, D. Perito, C. Fabry, and J. Mitchell. 2011. The Failure of Noise-Based Noncontinuous Audio Captchas. In 2011 IEEE Symposium on Security and Privacy. 19–31.
https://doi.org/10.1109/SP.2011.14

Elie Bursztein and Steven Bethard. 2009. Decaptcha breaking 75% of eBay audio CAPTCHAs.In proceedings of the USENIX Workshop on Offensive Technologies (WOOT'09).
https://www.usenix.org/legacy/events/woot09/tech/full_papers/bursztein.pdf

Reducing the Cost of Breaking Audio CAPTCHAs by Active and Semi-Supervised Learning by Malte Darnstadt and Hendrik Meutzner, Dorothea Kolossa https://www.ruhr-uni-bochum.de/lmi/darnstdt/captchas_icml2014.pdf

Assessing Threat Posed to Video CAPTCHA by OCR-Based Attacks by Alex Canter
https://www.cs.rit.edu/~dprl/old/files/AlexCanter_MScReport.pdf

A Three-Way Investigation of a Game-CAPTCHA: Automated Attacks, Relay Attacks and Usability Manar Mohamed, Niharika Sachdeva, Michael Georgescu, Song Gao, Nitesh Saxena†, Chengcui Zhang, Ponnurangam Ku-maraguru, Paul C. van Oorschot, Wei-Bang

Chen https://cpb-us-w2.wpmucdn.com/sites.uab.edu/dist/3/66/files/2020/01/2014-ACMCCS-1.pdf

Filip Vitas. 2019. How to bypass “slider CAPTCHA” with JS and Puppeteer. https://medium.com/@filipvitas/how-tobypass-slider-captcha-with-js-and-puppeteer-cd5e28105e3c

Carlos Javier Hernandez-Castro and Arturo Ribagorda. 2010. Pitfalls in CAPTCHA design and implementation: The Math CAPTCHA, a case study. Computers \& Security 29, 1 (2010), 141 – 157. https://doi.org/10.1016/j.cose.2009.06.006

2016. Sweet CAPTCHA solver. https://github.com/drdre1/Adultddl-Sweet-Captcha-Solver

Suphannee Sivakorn, Jason Polakis, and Angelos D. Keromytis. 2016. I’m not a human : Breaking the Google reCAPTCHA. In BlackHat 2016. https://www.fotolia.ir/cdn/2017/12/asia-16-Sivakorn-Im-Not-a-Human-Breaking-the-Google-reCAPTCHA-wp.pdf

Ismail Akrout, Amal Feriani, and Mohamed Akrout. 2019. Hacking Google reCAPTCHA v3 using Reinforcement Learning. ArXiv abs/1903.01003 (2019). https://arxiv.org/pdf/1903.01003.pdf

Breaking e-Banking CAPTCHAs Shujun Li, Syed Amier Haider Shah, Muhammad Asad Usman Khan, Syed Ali Khayam, Ahmad-Reza Sadeghi and Roland Schmitz http://kops.uni-konstanz.de/bitstream/handle/123456789/6246/ACSAC2010_Full.pdf?sequence=1

https://www.kasada.io/captcha-is-obsolete/

https://www.perimeterx.com/resources/blog/2020/captchas-hard-for-humans-easy-for-bots/

https://saaspass.com/threats/prevent-captcha-attacks-with-two-factor-authentication/

https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8665729

Re: CAPTCHAs – Understanding CAPTCHA-Solving Services in an Economic Context Marti Motoyama, Kirill Levchenko, Chris Kanich, Damon McCoy, Geoffrey M. Voelker and Stefan Savage https://www.usenix.org/legacy/event/sec10/tech/full_papers/Motoyama.pdf
CAPTCHA Smuggling: Hijacking Web Browsing Sessions to Create CAPTCHA Farms http://seclab.nu/static/publications/sac2010captcha.pdf

Imperva. 2020. 2020 Bad Bot Report.

Audio CAPTCHA with a few cocktails: it’s so noisy I can’t hear you https://www.cl.cam.ac.uk/~is410/Papers/cocktails_draft.pdf

POSTER: I Can't Hear This Because I Am Human: A Novel Design of Audio CAPTCHA System https://www.researchgate.net/profile/Jusop-Choi/publication/325480906_POSTER_I_Can%27t_Hear_This_Because_I_Am_Human_A_Novel_Design_of_Audio_CAPTCHA_System/links/5f2b8d64a6fdcccc43ac8469/POSTER-I-Cant-Hear-This-Because-I-Am-Human-A-Novel-Design-of-Audio-CAPTCHA-System.pdf

Capture the Bot: Using Adversarial Examples to Improve CAPTCHA Robustness to Bot Attacks Dorjan Hitaj, Briland Hitaj, Sushil Jajodia, and Luigi V. Mancini https://arxiv.org/pdf/2010.16204.pdf