A SURVEY OF CURRENT RESEARCH ON CAPTCHA

Walid Khalifa Abdullah Hasan

Aljabel AlGharbi University,Faculty of Eng-Jadu ,EE dept, Gharyan, Libya

ABSTRACT

The internet has been playing an increasingly important role in our daily life, with the availability of many web services such as email and search engines. However, these are often threatened by attacks from computer programs such as bots. To address this problem, CAPTCHA (Completely Automated Public Turing Test to Tell Computers and Humans Apart) was developed to distinguish between computer programs and human users. Although this mechanism offers good security and limits automatic registration to web services, some CAPTCHAs have several weaknesses which allow hackers to infiltrate the mechanism of the CAPTCHA. This paper examines recent research on various CAPTCHA methods and their categories. Moreover it discusses the weakness and strength of these types.

KEYWORDS

CAPTCHA (Completely Automated Public Turing test to tell Computers and Human Apart), Bot programs, Hackers, spammer, Scrapers ,search engines.

1.INTRODUCTION

The internet contributes profoundly to many aspects of human life such as communications, education, and online commercial activities etc. Some web services have online registration where the users provide information in order to connect and use services such as email in Yahoo, Gmail and Hotmail. However, many programs have been developed by hackers which automatically complete website registration pages with incorrect information which can cause traffic congestion, minimizing the performance of the system and in some cases, even causing it to fail, particularly where a website has a huge number of accounts.

Therefore, researchers developed a mechanism to distinguish between human users and computer programs in the case of online registration. The standard mechanism used currently to address this issue is CAPTCHA (Completely Automated Public Turing Test to Tell Computers and Humans Apart). The concept of CAPTCHA is based on the ability of humans to do certain tasks which computer programs cannot, such as asking users to type a distorted text image or choose a particular picture from many displayed pictures.

In recent years, many types of CAPTCHAs have been developed. Some are based on Optical Character Recognition (OCR) such as text CAPTCHA, whereas others are based on Non-Optical Character Recognition (Non-OCR) which uses multimedia, such as voice and video.

Some of these types of CAPTCHAs have been broken by new bot programs. For example, a text CAPTCHA can be broken by using the mechanism of segmentation letters. This paper will describe the new types of CAPTCHAs proposed in recent published papers, describe their classifications, comparing between CAPTCHA based on weakness and strength of them.

2.BACKGROUND

2.1. Definition of CAPTCHA

"A CAPTCHA is a program that protects websites against bots by generating and grading tests that humans can pass but current computer programs cannot. For example, humans can read distorted text as the one shown below in figure 1 but current computers cannot" [1].

Figure 1. Concept of CAPTCHA.

2.2.Application of CAPTCHA

1. Protecting Website Registration: CAPTCHAs are utilized to protect several free E-mail services such as Yahoo, Gmail, and Hotmail from Bot programs that would register thousands of email accounts every minute using automated script [1],[2].
2. Protecting Email Addresses from Scrapers: This can be done by hiding a user's email address from Web scrapers by asking the user to solve a CAPTCHA before displaying his or her email address [1].
3. Online Polls: CAPTCHAs are also utilized to restrain Web crawlers and bots from engaging in online polling by asking the user who wants to vote to solve a CAPTCHA before the vote submission. However, this process cannot prevent users from voting many times [2][1].
4. Preventing Dictionary Attacks: This is to restrain computer programs from being able to repeat through the entire space of password by asking the user to solve a CAPTCHA test after a number of unsuccessful logins. This mechanism is better than locking an account after a certain number of unsuccessful logins. [1][2].
5. Search Engine Bots: CAPTCHAs can be utilized by administrators to stop search engines from indexing to prevent others from downloading or reading these sites because sometimes they contain private information [1][2].

2.3. History of CAPTCHAs

- 1996 Moni Noar suggested the use of an Automated Turing Test to distinguish between human users and bots. [3].
- 1997 Andrei Broder et al. developed a mechanism to distinguish between human users and computer programs and also in the same year, the Altavista website used this method to block bot programs from entering by displaying a distorted English word to the user and asking the user to copy it [4]
- 2002 Broder announced that a CAPTCHA system had been in place for more than year which had minimized the number of spam advertising URLs by more than 95% [5].

- 2000 term CAPTCHA was coined by the team led by Manuel Blum and Luis von Ahn at Carnegie Mellon University [6][7].
- 2003 Barid and Monica Chew from California designed the Baffle text CAPTCHA [7].
- 2004 the Yahoo website utilized a simple version of the EZ-Gimpy method [4].

3. CLASSIFICATION OF CAPTCHAS

The meaning of CAPTCHA introduces a challenge response test to the users even though human or bot programs. The CAPTCHAs can be classified into five types depend on what is distorted that is whether characters, digits, or images .These types are given below:

1. CAPTCHAs based on text.

2. CAPTCHAs based on image.

3. CAPTCHAs based on audio.

4. CAPTCHAs based on video.

5. CAPTCHAs based on puzzle.[8].

3.1.CAPTCHA Based on text

Text based is simple to implement where a sequence of letters and digits present to the user with adding some modifications to the characters such as noise, scattering, rotating or making characters in the form of 3D these modifications added to prevent bot programs from reading the actual characters. A number of methods based on text are presented as below [8]:.

3.1.1.Gimpy CAPTCHA

The technique underpinning this method is to select a sequence of characters and present them as distorted and corrupted images by adding black and white lines and making non-linear modifications and asking the user to type the characters correctly. Gimpy was improved in collaboration with Yahoo and was implemented to prevent spammers from posting classified ads in chat rooms and writing scripts to create free e-mail accounts. Figure 2 is an example of the Gimpy method [9].
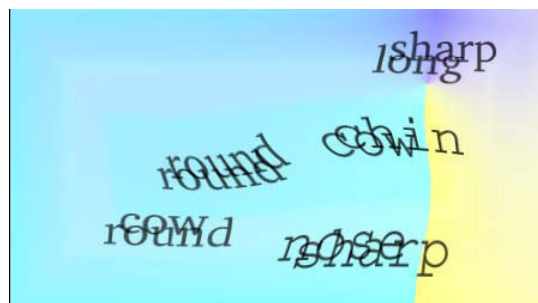

Figure 2. Example of a Gimpy CAPTCHA.

3.1.2.EZ-Gimpy

This is a simple version of the Gimpy CAPTHCA which is easy to identify compared to a Gimpy CAPTCHA. This method is used in chat rooms in Yahoo and is also used for authentication purposes when signing up. Figure 3.shows an example of EZ-Gimpy[9].
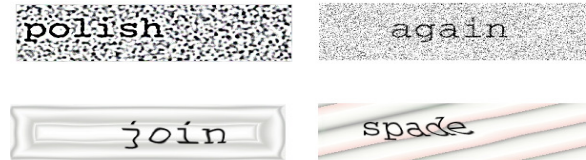

Figure 03. Example of EZ-Gimpy

3.1.3.PessimalPrint

There is not much difference between this type of CATCHPA and Gimpy and Baffle text. The CAPCTHA here utilizes a model of a document image degradation that is nearly ten aspects of physical of machine printing and images of text. This model includes spatial sampling rate and error, symbol size, blurring and thresholding [9].Figure 4 shows an example of Pessimal Print

Figure 4. PessimalPrint Method

3.1.4.Baffle Text

Baffle text was developed by Monica Chew (UC Berkeley) and Henry Baird (PARC) at the Palo Alto Research Centre. In this type of CAPTCHA, words which are not part of the British vocabulary is shown to the user , These words are changed to their image such as printing out the image and scanning it back in, or applying the threshold technique which transfers the image from color to black and white and back again. This changes the gray level and adds noise randomly to the image. The idea behind Baffle text is to minimize the small dictionary problem by using nonsense words where a human can use inference to solve the problem, but computer programs cannot [9][4]. Figure 5 shows an example of a Baffle Text CAPTCHA.

Figure 5.Example of a Baffle Text CAPTCHA

3.1.5.Scatter Type Method

This type of CAPTCHA involves the segmentation of letters. Letters are modified by cutting each word into displaced pieces. The existing methods for the recognition of letters cannot separate

letters easily because each letter in this method is segmented into numerous small pieces. Furthermore, the letters are chosen randomly so vocabulary cannot be utilized to predict the word. Figure 6 shows an example of the Scatter type method [9][4].

Figure 6. Example of a Scatter type CAPTCHA

3.1.6.Clickable CAPTCHA

Figure 7. Example of Clickable CAPTCHA

The user must click on the three cells which contain English words to pass the test but if the user clicks on words which are not part of the English language, the solution will be invalid. In this method, many variations are possible, for example, different grid sizes, any textual CAPTCHA may be used as an alternative for Google CAPTCHAs, and any binary textual classification problem can be used instead of recognition of English words [10].

This method requires the user to be proficient in the English language and also, it needs a mechanism to recognise that the user is using a mobile or small device with limited resources.

3.1.7.Strangeness in Sentences CAPTCHA (SS-CAPTCHA)

Yamamoto, Tygar & Nishigaki designed a Strangeness in Sentences CAPTHCA [11]. This method is based on the ability of humans to distinguish natural sentences created by humans from machine-translated sentences. The computer shows the user many sentences which are a combination of natural and machine-translated sentences and asks the user to select the one which is created by a human. Figure 8 shows the concept of SS-CAPTCHA.
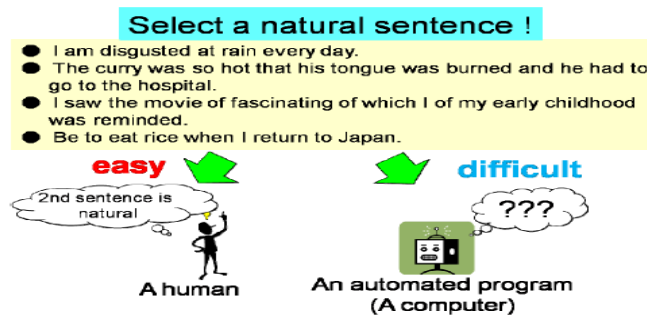

Figure 8.Concept of SS-CAPTCHA

Both natural sentences (NS) created by humans and the Non-Grammatical sentences (NG) created by a machine translation program are collected by this method.

NS are collected from paper media such as newspapers, magazines and books which is better than extracting them from the Internet because automated programs can easily search for sentences on the Internet and utilize them in the SS-CAPTCHA after recognizing them, whereas Non-Grammatical can be created by translating natural sentences from a non-mother-tongue into a mother-tongue language using machine-translation programs [11]

This approach requires the user to be proficient in the English language and is therefore not suitable for non-English speakers.[12].

3.1.8.3D CAPTCHA

Imsamai & Phimoltares improved the Text CAPTCHA by utilizing 3D characters [5].This method is based on the ability of humans to recognize a sequence of 3D characters easily whereas computer programs cannot. The CAPTCHA text consist of 6 alphanumeric characters with letters and numbers. Several adjustments have been made to the characters to increase the level of difficulty of breaking this type of CAPTCHA as follows:

1. 3D CAPTCHA with rotation: All characters are the same color but every character has a specific rotation angle on three axes, as shown in Figure 9 [13].

Figure 9. 3D CAPTCHA with rotation

2. 3D CAPTCHA with overlapping characters: In addition to being rotated as described in the previous method, the characters are also overlapped, as shown in Figure 10 [13].
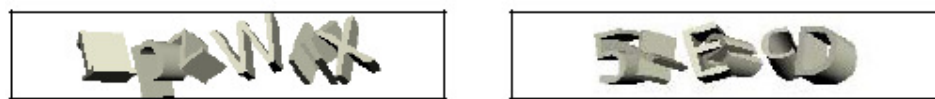

Figure 10. 3D CAPTCHA with overlap and rotation

3. 3D CAPTCHA with rotation and straight line:In addition to being rotated, a thick line is drawn through the characters, as shown in Figure 11 [13].

Figure 11. 3D CAPTCHA with rotation and straight line across a sequence

4. 3D CAPTCHA with distributed noise: In addition to being rotated, noise is added to the image, as shown in Figure 12 [13].

Figure 12. 3D CAPTCHA with distributed noise

5. 3D CAPTCHA with rotation and color background: In addition to being rotated, color is randomly added to the background and the text in the CAPTCHA, as shown in Figure 13 [13].

Figure 13. 3D CAPTCHA with rotation and color background

6. 3D CAPTCHA with rotation and character color variation: In addition to being rotated, each character is randomly coloured to attract the user and increase usability, as seen in Figure 14 [13].

Figure 14. 3D CAPTCHA with rotation and character color variation

7. 3D CAPTCHA with rotation and character scaling: In addition to being rotated, each character is scaled individually on each axis, as shown in Figure 15 [13].

Figure 15. 3D CAPTCHA with rotation and character scaling

8. 3D CAPTCHA with rotation and font variation: In addition to rotation, the font of the characters is changed to either serif or san-serif, as shown in Figure 16 [13].

Figure 16.3D CAPTCHA with rotation and font variation

9. 3D CAPTCHA with special character: In addition to rotation, a special character, such as %, @, $! is added, as shown in Figure 17 [13].

Figure 17. 3D CAPTCHA with special character

10. 3D CAPTCHA with background texture. In addition to rotation, the background layer is given a random texture which makes it difficult for computer programs to recognize and name each character, as shown in Figure 18 [13].

Figure 18.3D CAPTCHA with background texture

However, some types of 3D CAPTCHAs are confusing to users because some letters have similar shapes after rotation and overlapping [13].

3.2.CAPTCHA based on image

Image based CAPTCHAs are based on recognizing a specific image from similarity images and sometimes are mixed with some words. It is difficult for Bot programs to identify the graphics where human users can solve that. Various methods which are image based are shown as below [8]:

3.2.1.Bongo

This type of CATCHA requires the user to solve a visual pattern recognition problem. Two series of blocks are displayed, one on the left and one on the right. The blocks on the left differ from those on the right, and the user must find the characteristic that sets the two series apart .Figure 19 shows an example of the Bongo method [9].
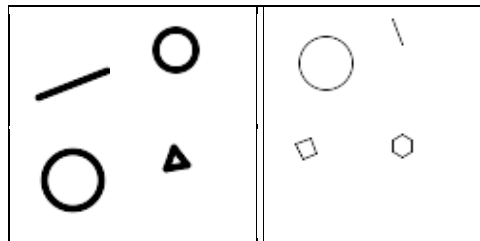

Figure 19. Bongo Method

 In this case, the difference between the left and right series is that the blocks on the left are drawn with thick lines, while those on the right are drawn with thin lines. After determining the characteristic which separates the two series, the user is shown four single blocks and is asked to decide on which side they belong. The user passes the test if they correctly determine the side to which all four blocks belong [1].

3.2.2.Pix CAPTCHA

The Pix CAPTCHA uses a large database of pictures and animated images of everyday objects such as dogs, flowers etc. The user is shown four different pictures of the same object and is required to type a word to indicate the object or concept to which all images belong [9] Figure 20 shows an example of the Pix method.
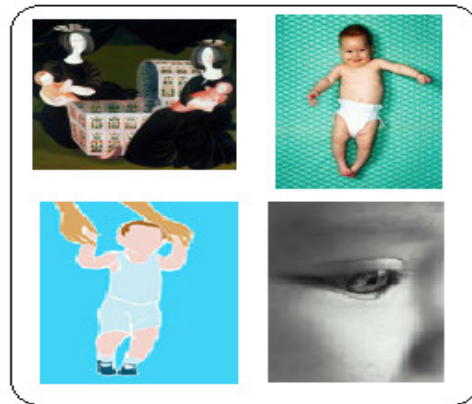
Figure 20. Pix method

3.2.3.Hand-Written CAPTCHA Method

This method uses a database of handwritten names of American cities, gathered from postal letters. An image of a handwritten city name is shown to the user who is then required to type the name of the city. The quality of the picture of the word which is shown to the user is poor and some are quite difficult for human users to read [4].

3.2.4.Implicit CAPATCHA Method

In this method, the user will chose a specific picture by a simple click. For example, the image of a mountain is displayed to the user and the user is asked to click on its top, as shown Figure 21. Conversely, a number of words may be shown and the user is asked to click on a specific word. Although this method is costlier, it is easier for users [4].

Figure 21. Implicit CAPTHCA method

3.2.5.Drawing CAPTCHA Method

This method is used in devices using style such as PDA (Personal Digital Assistance). In this method, a screen with a noisy background is prepared and sketch many dots on it then the user is asked to connect certain dots to each other, as shown in Figure 22. Humans can complete this task easily whereas computers cannot. Another merit of this method is that the user does not require to have ability of using a specific language [4].
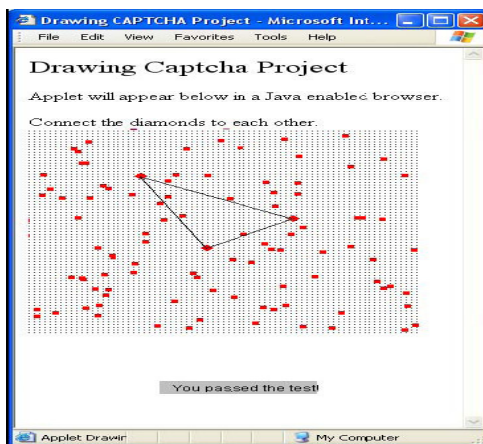
Figure 22. Example of a Drawing CAPTCHA

3.2.6.Four-Panel Cartoon CAPTCHA

Yamamoto, Suzuki & Nishigaki proposed a Four-Panel Cartoon CAPTCHA [14]. Their approach is based on a human's ability to understand humor. A four-panel cartoon is presented with the four panels rearranged randomly and the user is required to arrange them in correct order. Humans can understand the meaning of the pictures whereas computers cannot. However; this method requires a higher level of security because if there are 4 panels, then there are 24 possible combinations, so computers and software such as Malware could respond with correct answer. Figure 23 shows an example of a Four-Panel Cartoon CAPTCHA [14].

1st Panel	フーン	Hmmm…
1st panel	キノコ狩りにいったのしかったよー	I went on a mushroom hunt. It was fun.
2nd panel	庭に白いきもちわるいのがはえてるんだなんとかして	There are some gross white things growing in the yard. Go do something about them.
3rd panel	ワーイ	Yay!
3rd panel	いいよ	Sure.
4th panel	ボクもキノコ狩りにいきたいー	I want to go on a mushroom hunt too!
4th panel	おじいちゃんにたのんでみたら？	Why don't you go ask your grandpa?

Figure 23. Example of a Four-Panel CAPTCHA

3.2.7.Question-based CAPTCHA

M.Shirali-shahreza & M.Shirali-shahreza developed a Question-based CAPTCHA [18], merging OCR-based and Non-OCR-based methods. A simple mathematical problem is shown to the user in the form of images and the user is asked to answer this question. The images are selected randomly from a database of images and can be changed. For example, "*There are 5 cats, 3 apples and 4 dogs. How many pets are there on the table?*" [15]. Figure 24 shows an example of a Question-based CAPTCHA.
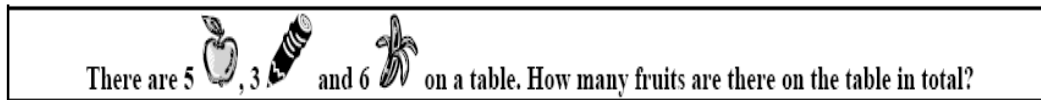
Figure 24.Example of a Question-based CAPTCHA

Also, the CAPTCHA can place an image in the question instead of text as in the following example in Figure 25.
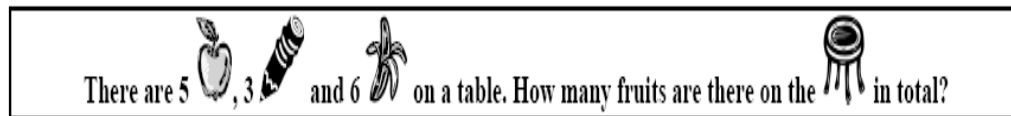

Figure 25.Example of using images in the question.

Humans can answer this question easily whereas it is very difficult for computers to recognize the phrases and shapes which are shown and also understand the question. This method has many advantages as follows:

1. This method is easy as the user only needs to type one number for the answer, hence it is also very time effective for the user.

2. In this method, the keyboard is not important because the user only has to enter a number so this method is useful on devices which do not have keyboard or on devices where it is difficult to utilize a keyboard, such as mobiles and Pocket PCs.

This method does not require any processing to be executed by the client and it can be used on small devices and on devices with limited resources [15].

3.2.8.CAPTCHA Systems for Nintendo

Nintendo DS is a well-known game console, invented in 2004. Today, it has more functions such as Internet browsing, education, playing audio and video and chat. The Nintendo DS Browser is used to connect Nintendo DS to the Internet but it does not support certain functions such as Adobe Flash. So, readily available CAPTCHA methods cannot be used for the Nintendo DS. M.Shirali-shahreza & S.-shahreza. proposed a CAPTCHA method for the Nintendo DS [17]. The mechanism of this method is based on the Online Collage CAPTCHA but is modified to make it suitable for this device which has browser limitations, namely to limit the number of objects to four because the Nintendo DS browser has a small memory of around 10 MB so it does not have the ability to load many images and large web pages. Also, small pictures are chosen to minimize size and it does not rotate pictures because the Nintendo DS screen is small which would make it difficult for humans to recognize small rotated pictures. [17].
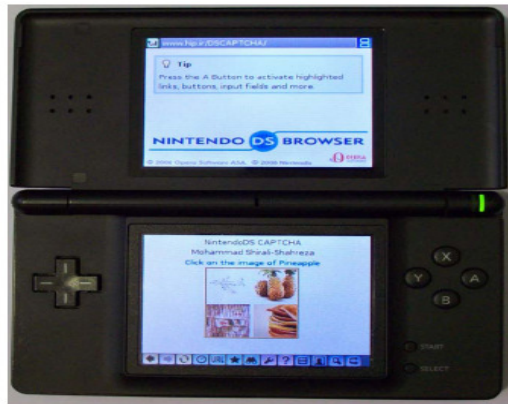
Figure 26 A screen shot of the Nintendo Ds CAPTCHA method.

This method requires the user to be proficient in the English language and also it needs a mechanism to recognise the user is using a Nintendo device.

3.2.9. CAPTCHAs for Children

This method is suitable for children and is based on finding the picture of an object from other objects. This approach was developed by M.Shirali-shahreza and S.Shirali-shahreza [18]. The first step in this method is to prepare a bank of names of various objects, for example, cars, countries, flags, animals selected randomly by the program. Pictures of these objects are downloaded from the Internet using websites such as Yahoo and the size of these pictures is chosen as medium. Then each picture is rotated and the program displays the pictures randomly to the user. The computer asks the user click on one particular picture using the Text-To-Speech system. For example, the program shows the user four pictures: a car, rabbit, children, and a bird and asks the user to click on the picture of the rabbit by saying this word. If the user clicks on the picture of the rabbit, it is concluded that the user is a human and not a computer program [18].Figure 27 shows an example of a CAPTCHA for children.

Figure 27. Example of a Children's CAPTCHA

There are many advantages of this type of CAPTCHA

1. This method does not need skills in reading and writing because the question is spoken to the user who is then required to simply click on the correct picture, hence it is easy for children of all ages.

2. This type of CAPTCHA is not time consuming to complete and does not frustrate users.

3. This method does not require a keyboard or any processing on the client side so it is suitable for devices which have limited resources.

4. In this method, there is no need to create a bank of pictures because the Internet is used to choose a picture so the cost is reduced [18].

This type of CAPTCHA requires the child to have English language skills and it also needs a mechanism to recognize that the user is child.

3.2.10.Collage CAPTCHA

M.Shirali-shahreza and S.Shirali-shahreza proposed the Collage CAPTCHA [19], a simple method which uses a bank of pictures of objects such as flags, animals, furniture etc and randomly selects six pictures and rotates them slightly, but does not overlap them. These pictures are then shown to the user who is required to click on one of them, as shown in Figure 28 [19].

Figure 28.Collage CAPTCHA

3.2.11.Advanced Collage CAPTCHA

Shirali-shahreza & Shirali-shahreza improved the Collage CAPTCHA [20] to maximize its resistance to attack, as follows:

After selecting one of the displayed pictures, the program chooses other objects, rotates them slightly and places them on the right side of the screen. One of the additional pictures is the same as the goal picture. The user is required to click on the goal picture and also on the same picture on the right side of the screen [20].

3.2.12.Online Collage CAPTCHA

Shirali-shahreza and Shirali-shahreza further improved the Collage CAPTCHA [21]. This method is the same as the previously described Collage CAPTCHA method, but the pictures are downloaded from the Internet which is able to provide a large bank of pictures and thereby reduce cost. Figure 29 show an example of this method [21].

Figure 29. Online Collage CAPTCHA

3.2.13.Multi-Modal CAPTCHA

Almazyad, Ahmad, Kouchay utilized the Multi-modal method to create a CAPTCHA [22], by combining Text-Based and Image-Based schemes where thousands of images (fruit, vegetables, animals) are gathered from search websites such as Google and Yahoo, and then this large set of images and text labels is stored in the database. The test is in the form of an image with four text labels and the user is required to write the correct answer in the cursive text label as shown in Figure 30.
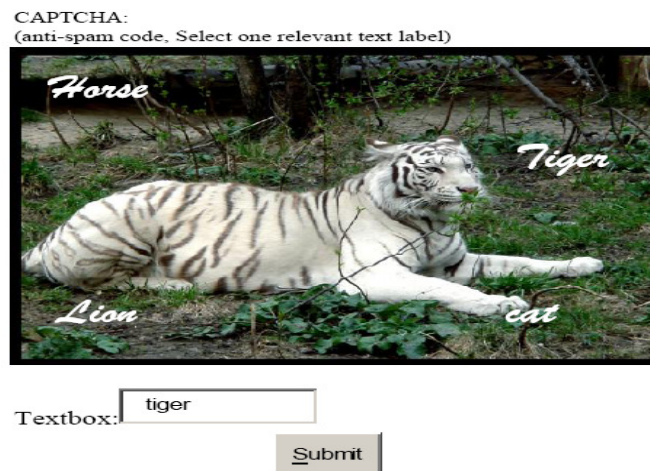

Figure 30. Example of a Multi-Modal CAPTCHA

This method has a very high pass level as the images used in this type of CAPTCHA are very simple so users who have a visual disorder can easily recognize them and solve this test [22].

However, this approach requires English language skills.

3.3.CAPTCHA based on audio

CAPTCHA based audio represent in introducing words to the user in the form of voice, so the user is requested to type what he/she has listened to solve the CAPTCHA. The advantage of the type is that it can be used for visually impaired users. Below is a discussion of some of the mechanisms [8].

3.3.1.Text –To Speech conversation method

In this method, a voice recording of a word or a number is presented to the user with some distortion and noise added and the user is required to type the word or number. Google uses this method for blind and low vision users for many Google services such as Gmail [9][4].

3.3.2.Spoken CAPTHCAs for Blind users

This CAPTCHA method was designed by Shirali-shahreza, Abolhassani, Sameti, & Shirali-shahreza for blind users who cannot respond to visual patterns [23].The overall structure of this method is shown in Figure 31. First, the system chooses one of the words from the system's dictionary and transmits this to the Text-To-Speech module which creates an audio file and asks the user to say the word. The user's reply is recorded and sent as a speech file to two modules on the server: one is Speech Recognition and Human vs. to recognize the word spoken by the user; the other is Computer Analysis which analyses the user's reply to decide whether this is a human voice or a computer program such as a Text-to-Speech system. The results from these two modules are then analysed to make a final decision. If the voice is from the user and also provides the correct answer, the user will be given access to the website, otherwise the system will send the user another test.[23].
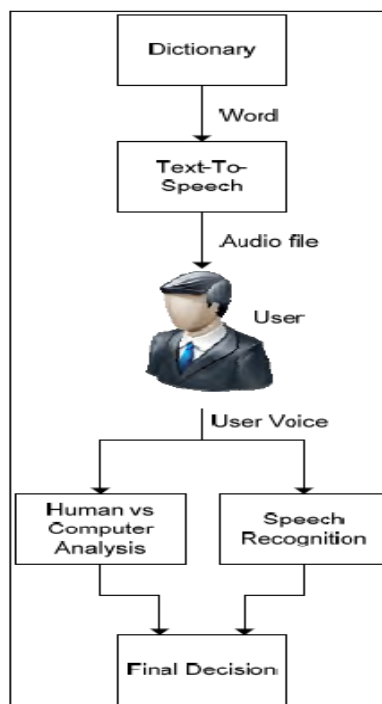

Figure 31.Structure of Spoken CAPTCHA

3.4.CAPTCHA based on video

Video CAPTCHA is rarely used in CAPTCHA system as it requires image download through the internet. One of the mechanisms is presented below [8]:

3.4.1.Motion CAPTCHAs

This method is improved by M. Shirali-shahreza. & S. Shirali-shahreza [24].A short movie clip of a person performing some type of action is shown to the user, after which a list of sentences describing various actions is shown to the user and the user is required to select the correct description .If the user selects the right answer, then the user will be allowed access to the website, otherwise the system will offer a new test to the user for a second chance. Figure 32 shows an example of this method [24].
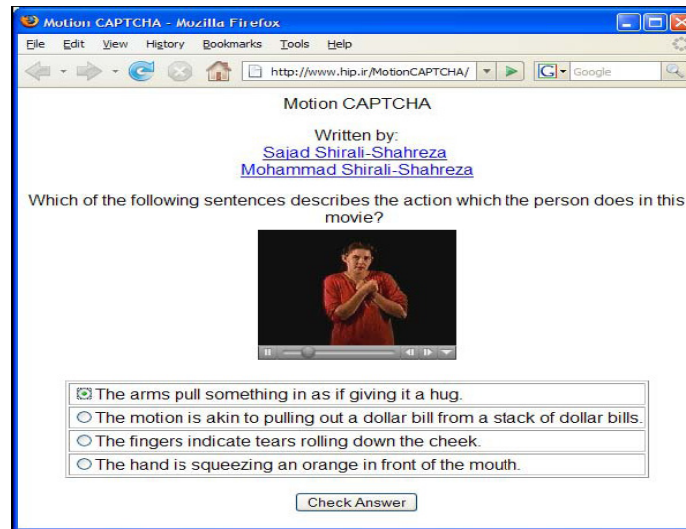

Figure 32.Example of a Motion CAPTCHA

Although this method is simple and easy, it requires ability in the English language.

3.5.CAPTCHA based on puzzle

In puzzle CAPTCHAs, the user are required solve a riddle which represent on introducing chunks of picture and asking user to combine the chunks or identify a specific part of image. Some of the methods in this category are given as below:[8].

3.5.1.Multiple SEIMCHA

A semantic image CAPTCHA has many levels of difficulty, from choosing a specific object in an image to identifying the upright orientation of an image which is difficult for computer programs. The simple idea behind this CAPTCHA was first introduced by Gossweiler et al. in the paper "What's up CAPTCHA" where a series of randomly rotated images were displayed and the user was required to rotate the image until it was in an upright position. Figure 33 shows a screenshot of this CAPTCHA.
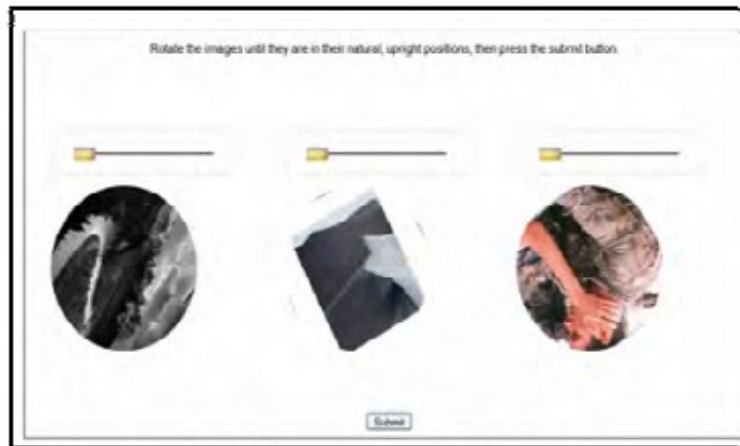
Figure 33.What's up CAPTCHA interface

Following this method, Ross et al. presented a new CAPTCHA which was based on an upright orientation of line drawings rendered from a 3D model called Sketcha. A large number of images from diverse angles were downloaded from the Google 3D Warehouse. Figure 34 shows the Sketcha method where the user must click on the image to rotate it until it is upright.
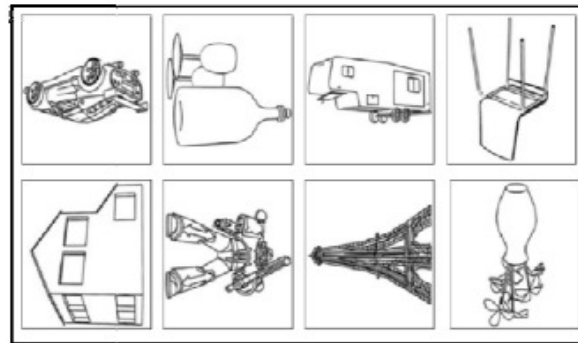

Figure 34. Sketcha method

Mehrnejad, Bafghi, Harati & Toreini used the concept of upright orientation and geometric transformation in Multiple SEIMCHA [25]. The structure of this type of CAPTCHA is as follows: firstly, a set of 3D images such as spheres and cones are created by transforming and warping the original image. Then, the image is modified using 2D projection. To pass this test, the user must click on the upright orientation of the final image. Figure 35 shows an example of the Multiple SEIMCHA.[25].
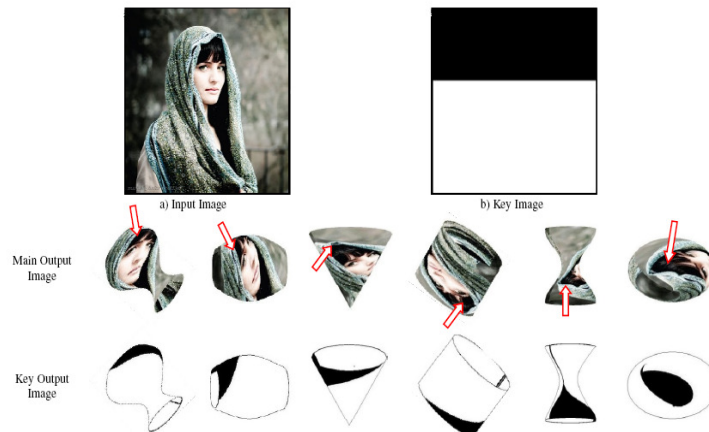
Figure 35.Example of Multiple SEIMCHA

3.5.2.A Novel Image-Based CAPTCHA Using A Jigsaw Puzzle

This method is created by presenting the user a jigsaw puzzle with two dislocated puzzle pieces from a puzzle database. All the jigsaw puzzle pieces are of rectangular shape. Some edge handling has been done to each jigsaw piece to rein edge detection. Figure 36 shows a simple 3*3 challenge and Figure 37 shows the completed puzzle.

Figure 36.Low level challenge   Figure 37.Completed puzzle

This type of CAPTCHA only uses images so it does not depend on language. Humans can easily and quickly solve this CAPTCHCA whereas it is very difficult for computers to pass this test [26].

4.COMPARISON BETWEEN CAPTCHAS TYPES

4.1.Strength and weakness of text based CAPTCHA

Text based CAPTCHA is the simplest type of CAPTCHA where it is the first type which has been invented and implemented in Email services and search engine. Text based CAPTCHA consists of English letters and numbers .these characters are limited so the bot programs and hackers can solve the text CAPTCHA by designing programs which scan the text CAPTCHA and typing it in the specific place. This problem is solved by doing some modification on characters such as adding some noise or rotate and scatter letters or present characters as corrupted and distorted letters or introduce characters as 3D .these modifications cause some problems to the

user when he identify the correct characters because some characters have similar shapes after making modifications.in addition,. Another issue in text based CAPTCHA is presented by English language where some users can't understand some types of text based CAPTCHA such as clickable CAPTCHA and Strangeness in Sentences CAPTCHA.[8],[27].

4.2.Strength and weakness of CAPTCHA based on image

In this type, some similarity images are introduced to the user to select the suitable image depend on the question under the CAPTCHA .Although this type is simple there are some problems can face the user when he is trying to solve image based CAPTCHA:

1. Some users who have low vision or learning disability will meet some issues when they are attempting to solve this CAPTCHA

2. Probability of break the CAPTCHA by bot programs will increase if the number of choice is decreased so it is better to create more options in CAPTCHA to make it strong however this mechanism will consume the database.

3. This CAPTCHA is available just by English language therefore English speakers and some others who have English vocabulary knowledge will solve the CAPTCHA only.[8],[27].

4.3.Strength and weakness of CAPTCHA based on audio

This type of CAPTCHA is designed for people have visual problem where the recorded words are introduced to the user and the user should type the word which he/she has listen.

1. Although audio CAPTCHA is available for visually disabled users, there are some issues may face the users:

2. The added noise to the recorded words is to make the CAPTCHA stronger and product it from breaking by Bot programs, can confuse the user and leading to fault answer and may

3. Audio CAPTCHA is introduced by English language therefor just users with English ability can solve this type.

4. In English language there are some letters have similar sound such as J &G, C & K .this can cause confusing to the user. [8], [27].

4.4.Strength and weakness of CAPTCHA based on video

CAPTCHA based on video represents in introducing a short movie contains person representing some kind of action and the user must select correct description from the list.

The size of video in this method is large so the users will face issue when they downloading it from the internet .this problem can lead users to leave the website or Email which they was trying to utilize it. Another issue can affect some users who do not have English language ability because video CAPTCHA is available by English language only.[8].

4.5.Strength and weakness of CAPTCHA based on puzzle

Puzzle CAPTCHA represents on introducing parts of images and the user should to merge the parts or identify a particular part of image. This mechanism consume more time to solve so user

may be bored and leave the website. Furthermore, users who have low vision will face image identification issue. [8].

5.CONCLUSIONS

CAPTCHA plays important role in World Wide Web security where it prevents Bot programs and Hackers from abusing online services. This paper has presented concepts and history of CAPTCHAs, and discussed their applications. This paper describing classification of current CAPTCHA methods based on text, images, voice, video and puzzle. In each classify there are many varied methods has introduced and discussed. In addition, we discussed the strength and weakness of each category. Finally, we propose to develop multi linguistics CAPTCHAs.

6.REFERENCES

[1]   Carnegie Mellon University, CAPTCHA: Telling Humans and Computers Apart Automatically. Available from: http://www.captcha.net/ [Accessed: April 22, 2015].

[2]   Pope, C. & Kaur, K. (2005), "Is it human or computer? Defending e-commerce with Captchas", IT Professional, vol. 7, no. 2, pp. 43-49.

[3]   Raj, S.B., Devassy, D. & Jagannivas, J. (2010), "A new architecture for the generation of picture based CAPTCHA", Electronics Computer Technology (ICECT), 2011 3rd International Conference on, Kanyakumari, 2011, vol. 6, no. pp. 67-71.

[4]   Shirali-shahreza, M. (2008), "Dynamic CAPTCHA ", Communications and Information Technologies, 2008. ISCIT 2008. International Symposium on. Lao.2008, vol., no. pp. 436-440.

[5]   Parc's   Captchas,   Parc's   Captchas.   Available   from: http://www2.parc.com/istl/projects/captcha/index.htm [Accessed: April 2, 2014,].

[6]   Almazyad, A.S., Ahmad, Y. & Kouchay, S.A. (2011), "Multi-Modal CAPTCHA: A User Verification Scheme", Information Science and Applications (ICISA), 2011 International Conference on,Jeju Island,2011., vol., no. pp. 1-7.

[7]   Cui, J.S., Zhang, W.Z.,Y .,Liang. Y., Xiao ,B., Mei., J.T., Zhang .,D. & Peng , W. (2010), "A 3-layer Dynamic CAPTCHA Implementation", Education Technology and Computer Science (ETCS), 2010 Second International Workshop on. Wuhan, vol. 1, no. pp. 23-26.

[8]   Singh Ved. & Pal preet. (2014), "Survey of different Types of CAPTCHA", international Journal of computer science and information technologies, 2014 vol.5, no. 2 pp. 2242-2245.

[9]   Chen Li, W.A., Wang, J. & Liu (2010), "Protection Through Multimedia CAPTCHAS",, vol., no. pp.

[10]   Chow, R., Golle, P. Jakobsson, M., Wang .L.& Wang , X.(2208), "Making CAPTCHAs clickable ", Conference HotMobile '08 Proceedings of the 9th workshop on Mobile computing systems and applications on. New York.2008, vol., no. pp.

[11]   Yamamoto, T., Tygar, J.D. & NishigakiM. (2010), "CAPTCHA Using Strangeness in Machine Translation", Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on, Perth, WA, 2010, vol., no. pp. 430-437.

[12]   Shirali-shahreza, S., Abolhassani, H., Sameti, H. & Shirali-shahreza, H. (2009), "Spoken CAPTCHA: A CAPTCHA system for blind users", Computing, Communication, Control, and Management, 2009. CCCM 2009. ISECS International Colloquium on, Sanya, 2009, vol. 1, no. pp. 221-224.

[13]   Imsamai, M. & Phimoltares, S. (2010), "3D CAPTCHA: A Next Generation of the CAPTCHA ", Information Science and Applications (ICISA), 2010 International Conference on, Seoul, 2010, vol., no. pp. 1 – 8. [14]

[14]   Yamamoto, T., Suzuki, T. & Nishigaki, M. (2011), "A Proposal of Four-Panel Cartoon CAPTCHA: The concept", Advanced Information Networking and Applications (AINA), 2011 IEEE International Conference on, Biopolis, 2011, vol., no. pp. 159-166.

[15]   Shirali-shahreza, M. & Shirali-shahreza, S. (2007), "Question-Based CAPTCHA", Conference on Computational Intelligence and Multimedia Applications, 2007. International Conference on,Sivakasi, Tamil Nadu.2007, vol. 4, no. pp. 54-58.

[16] Shirali-shahreza, M. & Shirali-shahreza, S. (2008), "A CAPTCHA system for Nintendo DS ", NetGames '08 Proceedings of the 7th ACM SIGCOMM Workshop on Network and System Support for Games on New York., vol., no. pp.

[17] Mehrnejad, M., Bafghi, A.G. Harati, A. &Toreini, E (2011), "Multiple SEIMCHA: Multiple semantic image CAPTCHA", Internet Technology and Secured Transactions (ICITST), 2011 International Conference on, Abu Dhabi, vol., no. pp. 196-201.

[18] Shirali-shahreza, S. & Shirali-shahreza, M. (2008), "CAPTCHA for children",System of Systems Engineering, 2008. SoSE '08. IEEE International Conference on, Singapore, 2008, vol., no. pp. 1-6.

[19] Shirali-shahreza, M. & Shirali-shahreza, S. (2007), "Collage CAPTCHA", Signal Processing and Its Applications, 2007. ISSPA 2007. 9th International Symposium on.Sharjah.2007, vol., no. pp. 1-4.

[20] Shirali-shahreza, M. & Shirali-shahreza, S. (2008), "Advanced Collage CAPTCHA", Information Technology: New Generations, 2008. ITNG 2008. Fifth International Conference on, Las Vegas, NV, 2008, vol., no. pp. 1234-1235.

[21] Shirali-shahreza, M. & Shirali-shahreza, S. (2007), "Online Collage CAPTCHA", Image Analysis for Multimedia Interactive Services, 2007. WIAMIS '07. Eighth International Workshop on.Santorini, vol., no. pp. 58-58.

[22] Almazyad, A.S., Ahmad, Y. & Kouchay, S.A. (2011), "Multi-Modal CAPTCHA: A User Verification Scheme", Information Science and Applications (ICISA), 2011 International Conference on,Jeju Island,2011., vol., no. pp. 1-7.

[23] Shirali-shahreza, S., Abolhassani, H., Sameti, H. & Shirali-shahreza, H. (2009), "Spoken CAPTCHA: A CAPTCHA system for blind users", Computing, Communication, Control, and Management, 2009. CCCM 2009. ISECS International Colloquium on, Sanya, 2009, vol. 1, no. pp. 221-224.

[24] Shirali-shahreza, M. & Shirali-shahreza, S. (2008), "Motion CAPTCHA", Human System Interactions, 2008 Conference on, Krakow, 2008, vol., no. pp. 142-1044.

[25] Mehrnejad, M., Bafghi, A.G. Harati, A. &Toreini, E (2011), "Multiple SEIMCHA: Multiple semantic image CAPTCHA", Internet Technology and Secured Transactions (ICITST), 2011 International Conference on, Abu Dhabi, vol., no. pp. 196-201.

[26] Gao, H., Dan, Y. Liu, H. Liu, X.& Wang, L (2010), "A Novel Image Based CAPTCHA Using Jigsaw Puzzle", Computational Science and Engineering (CSE), 2010 IEEE 13th International Conference on, Hong Kong,2010 , vol., no. pp. 351-356.

[27] Rizwan Rahman (2012), "Survey on CAPTCHA systems", Journal of Global Research in Computer Science, 2012 vol.3, no. 5 pp.

Author

Walid Hasan received his Bachelor of Electrical and Electronic Engineering degree from Algabal Algharbi University, Libya in 2006 and Masters of Telecommunication and Network Engineering from La Trobe University, Australia in 2012. He has also received a Postgraduate Certificate of Network from Swinburne University, Australia in 2013. Walid is currently a lecturer in the Department of Electrical and Electronic Engineering, Algabal Algharbi University. His research areas include Web security and Web Detection through Clustering Analysis.

Fei Liu was awarded the PhD degree by the Department of Computer Science & Information Technology, La Trobe University in 1998. She became a senior lecturer in the department since 2002. Before joining La Trobe as an academic staff, she was a lecturer in the School of Computer & Information Science, the University of South Australia, and the School of Computer Science & Information Technology, RMIT University. Fei also worked as a software engineer in Ericsson Australia. Her research interests include Text Mining, Logic Programming and Semantic Web.