IJFEAT

INTERNATIONAL JOURNAL FOR ENGINEERING APPLICATIONS AND TECHNOLOGY

TITLE: SURVEY PAPER ON TYPES OF CAPTCHA

Noopur R. Gawande, Shweta N. Kharode, Samiksha R. Tote

Student, Computer Science And Engineering, Jawaharlal Darda Institute of Engineering and Technology, Maharashtra, India, gawandenoopur@gmail.com

Student, Computer Science And Engineering, Jawaharlal Darda Institute of Engineering and Technology, Maharashtra, India, shwetakharode398@gmail.com

Student, Computer Science And Engineering, Jawaharlal Darda Institute of Engineering and Technology, Maharashtra, India, stote1998@gmail.com

Abstract

In our daily life, internet has been playing an important role. The access of various web services such as emails and search engines, use of internet is rapidly increasing .This growing use of internet often leads to certain attacks by the computer programs such as bots. To overcome such type of attacks CAPTCHA (Completely Automated Public Turing Test to Tell Computers and Humans Apart) was developed .CAPTCHA was first used by AltaVista in 1997 and reduce the spam by 95% for the small period of time. CAPTCHA is a type of challenge- response test. The CAPTCHA is the combination of various letters and numbers, bots would struggle to solve it. It is a distinguishable test to determine whether the user is human or not. The father of modern computing, Alan Turing proposed the test called Turing test. This test examines whether the machines can think like users are not. When it comes to the processing of visual data, computers lack the sophistication as compared to human beings. The human mind is able to perceive the patterns even when there exists none, this behaviour is called Pareidolia .Humans can pick up the patterns or images more easily than computers. The solution of CAPTCHA should not be stored in web page's meta data. The computer programs need a large database for different CAPTCHA questions. This can help in preventing the comments spam in blogs, online polling, preventing dictionary attacks, etc. Computer scientists have developed a new artificial-intelligence (AI) based programme which can easily decode text CAPTCHA schemes commonly used by websites to protect themselves from cyber attacks.

Index Terms: CAPTCHA, Application of CAPTCHA, Challenge test, Security, Image recognition.

1. INTRODUCTION

In security research and practice, authentication is the key area .For determining the access of user to the resource, the process of authentication is used. The conventional authentication approach for security is alphanumeric password which uses letters and some special characters .The another way for security is biometric where the password can be authenticated by human body .Nowadays ,advanced technologies make use of graphical password ,CAPTCHA .CAPTCHA(Completely Automated Public Turing Test To Tell Computers And Human Apart)is a test which is built by computer program which can pass by humans but cannot by bots .To protect the users access over the information from spam's ,the authentication system mechanism CAPTCHA is used .For designing a successful CAPTCHA system usability and robustness are the important aspects.

The CAPTCHA system possesses four important properties: -
1. Secure: By using any algorithm, machines can't solve the tests generated by the program.
2. Automated: The programs must be capable of generating and arranging the test.
3. Open: The algorithms and databases must be made public.
4. Usable: The tests must be solved by the humans as a user in a appropriate time.

1.1 Types of CAPTCHA system

The various types of CAPTCHA system are:

1. Text based CAPTCHA

2. Imaged based CAPTCHA

3. Audio based CAPTCHA

4. Video based CAPTCHA

5. Puzzle based CAPTCHA

1. Text based CAPTCHA: Text based CAPTCHA schemes ask users to recognize a string of distorted characters with/without an obfuscated background .Text based CAPTCHA is most widely used CAPTCHA. There are different types of text CAPTCHA's:

• Gimpy CAPTCHA: It is based on optical character recognition (OCR). The words are selected by the Gimpy from a dictionary and the process of misshaping is applied. The misshaped words are identified by the users.

Fig 1: Gimpy CAPTCHA

• Baffle text CAPTCHA: It was designed by Henry Baird. It is modified version of Gimpy. In this a random alphabets or characters are picked to form a pronounceable text.

Fig 2: Baffle Text CAPTCHA

• MSN text CAPTCHA: CAPTCHA makes the use of digits and upper case characters up to length 8.

Fig 3: MSN Text CAPTCHA

• Bar CAPTCHA: It consist of small bar that are used to specify text as well as adds noise in the background. Bots are unable to distinguish text and noise.

• Thread CAPTCHA: Thread appears as the handwritten text in a free form style using a long thread. It is similar to Bar CAPTCHA but in this bots cannot identify the thread which represents the text.
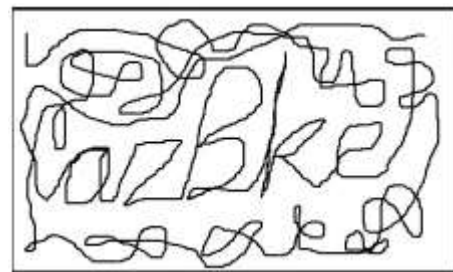

Fig 4: Thread CAPTCHA

• Transparent CAPTCHA: It is a technique in which the text is written in transparent font over an image. It makes difficult for the bots to identify the pixels that represent the exact text and the pixels that the other objects in the image.

Fig 5: Transparent CAPTCHA

2. Image Based CAPTCHA: It consists of collection of images integrated with same concept or object. The user is required to enter a concept which all the images are linked with. Some of the types of image based CAPTCHA are Bongo and Pix CAPTCHA, which uses different patterns that are identified by users but not by robots.

Fig 6: Image based CAPTCHA

- Bongo CAPTCHA: In Bongo, visual based pattern identification is provided for the user to solve .It contains right block and left block series. The user should identify how the right and left block series differ from each other.

- Pix CAPTCHA :

It uses large volume of databases containing images and photographs of the objects with label. The set of objects are picked at random from the database. Based on these objects the questions are asked. It is very easy to write a bots program for such question, as the data and the code of the CAPTCHA is displayed publicly which makes Pix to be not considered as CAPTCHA .However, one way to consider Pix as CAPTCHA is to distort the images randomly before displaying them to the user, so that the bot programs find difficulty in searching the database for undistorted images.

3. Audio based CAPTCHA:

It is for visually disabled a user which works on sound based system. It consist of downloadable audio files which the user should listen and submit the heard word .The program chooses a sequence of digits and words randomly and renders the words and number digits into sound clips and distorts it. The distorted sound clip is then presented to the user to enter the right word or number.

4. Video based CAPTCHA:

The system which uses a technique in which the video contains few random words. When the video is played the user has to submit the displayed words .It is not necessary for the user to wait until the video finishes. In this, three words are provided to the users. The users label must match the ground truth labels then only the test is said to be successful.

5. Puzzle Based CAPTCHA:

Puzzle based CAPTCHA can be classified as picture based puzzle and a mathematical puzzle. In a picture based puzzle, the picture is divided into segments and is shuffled. Each segment will have a segment number followed by the next segment. The user should arrange the puzzle appropriately, so that it could match the original one. Another one is mathematical puzzle which is 100% effective. On the basis of this the registration forms, login can be securely accessed.
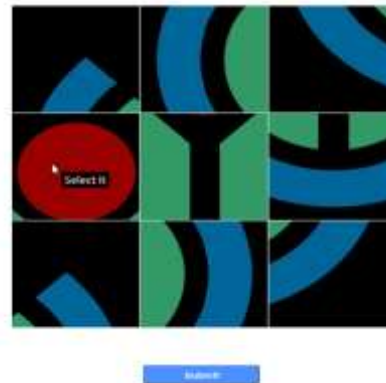
Fig 7: Puzzle based CAPTCHA

2. Advantages and Disadvantages of CAPTCHA
- Advantages:
1. Distinguishes between humans and bots.
2. Makes online polls more legitimate.
3. Makes online shopping safer.
4. Reduces spam and viruses.
5. Diminishes abuse of free mail account services.
6. May greatly enhance artificial intelligence.

- Disadvantages:
1. Time consuming to interpret.
2. Sometimes very difficult to read.
3. Technical difficulties with certain internet browsers.
4. They are not compatible with users having disabilities.

3. Application of CAPTCHA
- Protecting Website Registration - The Company which offers free registration services suffers from a type of attack called flooding attack.
- Preventing Dictionary attacks - Servers are hacked by entering words as a password from dictionary. After certain number of unsuccessful logins, system may ask the user to solve a CAPTCHA. Only humans can solve the challenge but bots can't.
- Online polling - For identifying whether the voter is human or not. The biased programs are prevented by automated

programs that ensures only humans should take part in polling.

- Smart Cards - To provide additional security to the smart cards while doing financial transactions, CAPTCHA is used other than username and password.

4. Comparison of different CAPTCHAs

Types of CAPTCHA	Security	Usability	
		Usable for visible impair people	Easy or difficult to use
Text-based	Good	No	Average
Image-based	Good	No	Easy
Video-based	Good	No	Difficult
Audio-based	Good	Yes	Difficult
Puzzle-based	Average	No	Difficult/Easy

Table: Comparison of CAPTCHAs

4. **CONCLUSION**

CAPTCHAs are an effective way to counter bots and spams. This paper includes different types of CAPTCHAs. The applications are varied from stopping bots to character recognition and pattern matching. There are certain issues with current implementations that represent challenges for future improvements. CAPTCHA though has many applications in varied fields, it is primarily used to distinguish between humans and bots.

REFERENCES

[1]. https://www.ukessays.com/essays/information-technology/applications-captcha-7274.php

[2]. https://www.v3.co.uk/v3-uk/news/3067828/new-ai-tool-can-easily-crack-text-captcha-schemes

[3]. https://www.ijecs.in/index.php/ijecs/article/download/1161/1059/

[4]. https://en.wikipedia.org/wiki/CAPTCHA

[5]. http://www.tjprc.org/publishpapers/1-27-1389261127-2.%20Cpatcha%20and%20its%20applications.full.pdf