

# Improved CAPTCHA Method

Rituraj Soni  
Assistant Professor CSE  
Engineering College Bikaner

Devendra Tiwari  
Assistant Professor CSE  
College of Engineering & Technology

## ABSTRACT

A CAPTCHA is a program that can tell whether its user is a human or a computer. CAPTCHAs are used by many websites to differentiate between the bots (automated program) and the human. The main motive behind using the CAPTCHA is to prevent abuse from "bots," or automated programs usually written to generate spam. The CAPTCHA basically uses images in distorted form that is difficult to read by the bots but they can be read by the human being easily. The websites which provide free services to the users some anonymous user can make false enrollment on the websites with the help of automated computer program. CAPTCHAs are used to prevent the false enrollment. One of the CAPTCHA methods is Collage CAPTCHA. In this method some shapes are shown with distortion and the user is asked to choose a specific object

In this paper we increase the resistance of this method to attacks. For this purpose, we show some objects on the left of the screen and some of the objects on the right side of the screen which contain the name corresponding to the images shown on the left side of the screen. Now we ask the user to choose object on the left side and a specific image (containing word) on the right side of the screen. The image on the right side of the screen contains the name of the image. After this user is asked to enter the name of the image in a text box shown below the two images. The user will be passed the test if he chooses the two similar objects correctly and then he/she has entered the name of the image correctly in the text box. In this method because the computer program should also recognize the similar object on the right side and need to enter the name of the image in text box, the possibility of passing the test by computer is more reduced.

## Keywords :

CAPTCHA (Completely Automated Public Turing test to tell Computers and Human Apart), Collage CAPTCHA, NON OCR CAPTCHA,

## 1. INTRODUCTION

With the development of the computer application in different field, internet has made a tremendous progress and become a special need in human life. It has application in a wide range of daily affairs including trade, education, daily purchases and dialogues take place with the use of Internet. One of the common actions in the Internet web sites, especially commercial and administrative ones, is to fill out registration forms for certain purposes. Unfortunately, there are some programs which

automatically fill out these forms with incorrect information to abuse the site. This wastes a large volume of the resources of the site. Moreover this type of activity on the sites leads to Denial of Service attack on certain sites. Due to such a problem the legitimate user cannot access the site in a proper manner.

CAPTCHA Completely Automated Public Turing test to tell Computers and Humans is a method used by many web sites to fight against computer-generated input. CAPTCHAs are challenge puzzles used to determine whether a user is human or not [1]. Intuitively, a CAPTCHA is a program that can generate and grade tests that most humans can pass but current computer programs cannot pass [2]. The goal of such a system is to ask simple question that a computer cannot pass, but a human can pass easily. CAPTCHA methods can be generally divided into two groups: OCR-based and Non-OCR-Based. Our method in this paper is a Non-OCR-Based method.

## 2. RELATED WORK

CAPTCHA was first introduced in 1997 when Andrei Broder devised the CAPTCHA method. In the same year, Altavista web site used this method to distinguish between computer programs and human user. In this method, a distorted English word was shown to the user and the user was asked to type it (Fig. 1). Distortion was needed so that OCR programs could not recognize the word



Fig: 1

As mention above CAPTCHA can be classified in two ways OCR based and non OCR based

### 2.1 OCR-Based CAPTCHA Methods

In OCR-based methods [7], the image of a word with distortion and various pictorial effects is shown to the user and he/she asked to type that word. Due to presence of various pictorial effects, the computer will encounter problems in the recognition of these words and only a human user can recognize the word. But these methods usually result in dissatisfaction of users. On the other hand, efforts have been made for attacking these methods [4]. Examples of these methods include Gimpy[2], Handwritten CAPTCHA, and Persian/Arabic CAPTCHA[6].

## 2.2. Non-OCR-Based CAPTCHA Methods

In contrast, we can point to Non-OCR-based methods which are easier to work with than OCR-based ones. Examples of these methods include PIX [8], Text-to-Speech method, and Drawing CAPTCHA.

Various methods have been proposed for breaking up the CAPTCHA. It can be done with the help of the segmentation and character recognition techniques. This can be done with the help of the Neural Network using divide and conquer approach. To provide defense against such techniques resistance method for CAPTCHA have also been proposed that includes Font tricks, Noise, Color model and Overlap[4].

## 3. The SUGGESTED METHOD

In this paper, we improved the resistance of Collage CAPTCHA method [2]. In this section, after describing the original Collage CAPTCHA method, we will describe the suggested algorithm.

### 3.1. Collage CAPTCHA Method

Collage CAPTCHA is a method for distinction between human users and computer programs through recognition and finding a picture of an object among some objects. The structure of the suggested method is as follows:

At first a bank of the pictures of objects is prepared. These pictures include objects, animals, different persons, flags of the countries, etc. The concerned program chooses some of these pictures (e.g. 6 pictures) randomly. Each of these pictures is rotated a little. Then all the pictures are shown on the screen. The pictures are putted in random places but the pictures don't overlap.

Computer chooses one of the shown pictures as a goal and asks the user to click on the picture of that object. If the user clicks on the goal picture, we can conclude that a human user has done the click and the user is not a computer program because a computer program should do two operations successfully in order to respond appropriately: First it must recognize the wanted object and has knowledge of its shape. Then it must find the place of the object on the screen.

### 3.2. The Improved Method

The improved method is designed in such a way to increase the resistance of collage CAPTCHA method to hackers attack. The structure is as follows

In this method we are displaying images on left side and right side of the screen. On the left side of the screen we have the images consisting of different objects like animals, different persons, objects like furniture, flags of countries etc. On the right side of the screen we have the image that consists of the name of the object shown on the left side of the screen.

Now the computer program asks the user to choose the picture on the left side of the screen and then he is asked to choose the corresponding image (containing name of the selected object). If the user selected both of them correctly, and after doing this user is allowed to enter the name of the image in the given text box, if user enters the name of the image correctly in the text box, then we can guess that user is human being, not a computer bot. More over the text box remained disabled until use selects both the

images correctly then only the text box is enabled and user can enter the name of the image in text box.

In this method computer requires four abilities to pass the test

1. To find out the shape of the concerned object
2. To find the concerned object on the screen
3. To find out the object containing the name of the "selected object" on the screen
4. To enter the name of the image in the text box

It is difficult for the computer to realize these tasks in correct order, only a human user can recognize and choose the concerned object.

This method can be implemented by Java programming language. The implementation is similar to original CAPTCHA method with some differences.

The CAPTCHA program select 8 images (objects) randomly that must be different from the previous images. The objects must be different from each other and should be placed on the left side of the screen. Then corresponding images containing the name of the objects already present on the left side should appear on the right side of the screen. The user is asked to select one of the image on the left side of the screen, say "a image of chair" and then he is asked to select the image containing the word "CHAIR" on the right side of the screen. If user has selected both the objects correctly and also entered the name of the image in the text box correctly, then he is allowed to do the concerned operations.

The example of the "chair" is shown in the fig 2.



Fig: 2

## 4. CONCLUSION

In this project we improve the Collage CAPTCHA method by presenting a method for increasing the resistance of it. Although we can increase the rate of its difficulty in order to improve its resistance against the attacks by applying other effects such as increasing the objects present in the screen and decreasing the distance between the objects, this way the test will become more difficult even for a human user in addition to computer programs.

Moreover we can also show the objects that can be overlapped in some fashion that will be difficult task for a computer program to identify the two objects correctly.

We can also implement this method in such a way that we can show the mirror image of the object on the right side of the screen

that would be a daunting task for a computer program to identify it correctly.

We can also implement this method on other devices such as mobile phone, PDA (Personal Digital Assistant), and the devices which have touch screens, because no keyboard is needed in this method and also there is no need for heavy processing. Besides its advantage of covering every age group, even children, we can also recommend it for disabled people. We can make this method specialized for specific websites for example we can use the specialized sport pictures (such as soccer players pictures) for sport news websites.

## 5. REFERENCES

- [1]. M.Shirali –Shahreza , S.Shirali Shaerza “Advanced Collage Captcha”, fifth international conference on information technology : New Generation.
- [2]. L.von Ahn M.Blum and J.Langford “Telling Humans and Computer Apart Automatically”,*Communication of the ACM* vol 47 , no2 feb 2004 pp 57 -60.
- [3]. R.Dhamija and J.D Tygar ,”Phish andHIPS :human second international Workshop on Human Interactive proofs ,(HIP 2005), PA , USA May 2005, pp 127 -141.
- [4]. Peter Gutmann ,David Naccache ,Charles C.Palmer,” CAPTCHAs: Human Vs Bots
- [5]. M.Shirali Shahreza ,”Highlighting Captcha “, Krakow , Poland May 25-27 2008
- [6]. M.H Shirali Shahreza” Persian/Arabic CAPTHCA” *IADIS international journal* vol 1 no.2 pp 63-75 october 2006
- [7]. K.Chelalpia and P.Simrad “Using Machine Learning to break Visual Human Interaction Proof(HIPS): in advance in Neural Processing System(NIPS) Vancouver CanadaMIT press December 2004
- [8]. M.Blum L.Von Ahn and J.Langford THE CAPTCHA project School of Computer Science Carnegie Mellon University ,November 2000