

Multilingual CAPTCHA

M. Hassan Shirali-Shahreza
Computer Engineering Department
Yazd University
Pejooohesh Street, Safa-ieh
P.O. Box: 89195-741
Yazd, IRAN
Phone: 0098-351-8247100
Fax: 0098-351-8210699
Email: hshirali@yazduni.ac.ir

Mohammad Shirali-Shahreza
Computer Science Department
Sharif University of Technology
Azadi Street
P.O. Box: 11365-9415
Tehran, IRAN
Phone: 0098-21-66165601
Fax: 0098-21-66005117
Email: shirali@cs.sharif.edu

Abstract- Currently, many daily human activities such as education, commerce and talks are carried out through the Internet. In cases such as the registering in websites, some hackers write programs to make automatic false enrolments which waste the resources of the website while this may even stop the entire website from working. Therefore, it is necessary to tell apart human users from computer programs which is known as CAPTCHA (Completely Automated Public Turing test to tell Computers and Human Apart). CAPTCHA methods are mainly based on the weak points of OCR (Optical Character Recognition) systems while using them are undesirable to human users.

In this paper a method has been presented for telling the human users and computer programs apart on the basis of choice of an object shown on the screen. The user interface of this method is multilingual. At the beginning of the test, the user chooses his/her native language. After that, all of the messages are shown in the selected language. The messages are translated using an online translator. While most of the available CAPTCHA methods assume that user knows English language, in this method, the user doesn't need to be familiar with English language.

In this method some objects are chosen randomly and the pictures about these topics are searched and downloaded from the Internet. Then all of the pictures are shown on the screen. After that, the user is asked to choose a specific object. The main advantage of this method is that non-English users can use it easily, even if they don't know English language. This method has been implemented by the PHP language.

Keywords: CAPTCHA (Completely Automated Public Turing test to tell Computers and Human Apart), Implicit CAPTCHA, Multilingual, OCR (Optical Character Recognition).

I. INTRODUCTION

Many aspects of human life have been affected by the expansion of the world-wide web, so that, in industrial countries, many daily affairs from daily shopping to education and commerce are all done on the Internet.

One of the common actions in most web sites, esp. commercial and administrative ones, is to fill out registration forms for certain purposes.

After filling out the forms by entering the required information, the individuals will be allowed to connect to that

web site to carry out certain jobs.

However, there are individuals who break the law by doing vandalistic acts such as writing programs to make automatic false registration in the web site. These programs automatically fill out a form with incorrect information to enroll in the site. This wastes a large volume of the resources of the site in favor of the profit-seeking programmers or reduces the performance of the system. Such attacks are known as "Denial of Services" or DoS.

Various methods have been presented in order to prevent such attacks, aiming at distinguishing human users from computer programs. The main characteristic of these methods should be their automaticity so as to be implemented only by using the computer because examination of a large bulk of registration on the Internet web sites by human forces requires a great deal of time and expense and in some cases, such as email services web sites, using human force for examining the registration forms is practically impossible. Therefore, it is necessary to use automatic systems to distinguish human users from computer programs.

In the discussions of artificial intelligence (AI), a test known as the Turing test is used for providing the intelligence of a computer [1]. In this test, a human person and a computer are put in two different rooms and a human interrogator in a third room asks them questions. If the interrogator cannot recognize which room the computer is in and which room the human, it is said that the computer has passed the Turing test.

A similar method to the Turing test should be used to distinguish human users from computer programs with the difference that the human interrogator should be replaced by a computer, which should ask questions to distinguish the human user from the computer programs. This method is called CAPTCHA (Completely Automated Public Turing test to tell Computers and Human Apart). Therefore the main focus of this method is on questions that the human user can easily answer but the present computer programs can hardly answer them.

One of the methods used for distinguishing human users from computer programs is the use of pictures of words. It is a method based on the weaknesses of current Optical Character Recognition (OCR) programs.

OCR programs are used for automatically reading the texts, but they have difficulty reading texts printed with a low quality or reading manuscripts and can only recognize high-quality typed texts that use common standard formats. So, this defect of the OCR programs can be taken advantage of by changing the picture of a word so that it can be recognized only by a human user but not by any OCR program. Section 2 will further elaborate on the methods used for this purpose.

The CAPTCHA methods are now used in big web sites such as Yahoo! and MSN. However, these methods usually disturb the users and also cannot be run on all systems because in mobile phones for example, it is difficult to type words. Besides these methods, in recent years, methods have been proposed for overcoming these methods and automatically recognizing such word images [2, 3].

But these methods are not comfortable for human users and most people use these methods reluctantly. In addition, non-English users have difficulties to pass these tests because the designers of these methods assume that the human user can read English texts easily.

Therefore limited methods in which there is no need to type and don't use OCR systems have been presented so far. Some of these methods are Implicit CAPTCHA [4], Text-to-Speech method [5], and Drawing CAPTCHA [6]. In the next section we will explain these methods.

In this paper a new method for distinction between human users and computer programs is proposed. This method is based on the showing of pictures. The picture includes some objects and the only thing that the users should do is to click on the wanted object. Therefore the computer requires two abilities to respond in this method:

1. To find out the shape of the concerned object.
2. To find the concerned object on the screen.

Considering the insufficient ability of the computers to carry out these tasks only a human user can recognize and choose the concerned object.

The user interface of this method is multilingual. So the user does not need any knowledge of English language and non-English users can use it comfortably.

In the third section we will discuss this method and its experimental results in full details. In the fourth section the advantages of this method will be discussed and in the last section we make the final conclusion.

II. RELATED WORKS

It was first in 1997 when Andrei Broder et al devised a system for distinguishing human users from computer programs. In the same year, Altavista web site used this method to tell computer programs and human apart. In this method, a distorted English word was shown to the user and the user was asked to type it (Figure 1).

Distortion was so that OCR programs could not recognize the word [7]. These systems were known as CAPTCHA systems and are now used in most well-known web sites such as Yahoo! and Microsoft. In this section we review some non-OCR based CAPTCHA methods.



Fig. 1. An Altavista CAPTCHA word [7]

A. Implicit CAPTCHA [4]

The common methods to tell human users and computer programs apart usually troubles the users because he has to read a text that is usually very difficult to read and then type it. However, in the Implicit CAPTCHA methods, the user has to make a simple click. For example, the picture of a mountain is shown to the user and he is asked to click on its top or a number of words are shown in an image and the user asked to click on a specific word. This method is an easier for the users although it is costlier.

This method has many similarities to our suggested method. We will compare these two methods in section III.

B. The PIX recognition method [8]

In this method, usual pictures (instead of pictures of words) are used to tell human users and computer programs apart. A library of pictures with different subjects is prepared for this method and a number of these pictures that have a similar subject is selected and shown to the user while asking the user to select the subject of the picture from among the subjects shown.

However, this method requires a large space for keeping the pictures and the library should also be very extensive, which requires large expenses.

C. Text-to-Speech conversion method [5]

In this method, instead of showing an image, a sound is played which has been obtained by converting text to speech by certain programs. The user must recognize and type the word.

Considering the many complexities of speech, it is very difficult for computer programs to recognize the played words. This method is also used by PayPal [9].

D. Drawing CAPTCHA [6]

This method is for devices using light pen such as PDA (Personal Digital Assistant). In this method, numerous dots are drawn on the screen with noisy background and the user is asked to connect certain dots to each other (Figure 2).

In view of the problems that computers face in recognizing the dots from the noise, only a human user can easily identify the special dots and connect them to each other.

Unlike the other CAPTCHA methods, the users of any language in any age group can use this program and run it on devices with more limited resources than the computer.

As we said earlier, there are also CAPTCHA methods which are using OCR like BaffleText CAPTCHA [8], ScatterType [10] and PessimPrint [11], which are different from our CAPTCHA method.

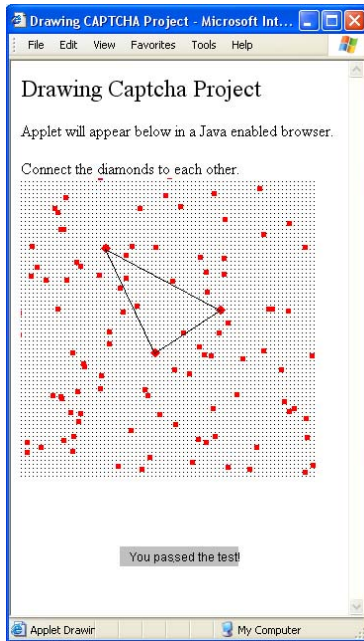


Fig. 2. Drawing CAPTCHA project [6]

Also there are some OCR-Based CAPTCHA methods for non-English languages. For example in Persian/Arabic Baffletext method, pictures of Persian or Arabic words are shown by adding some noise to their images [12].

III. THE PROPOSED METHOD

In this paper a method for distinction between human users and computer programs through recognition and finding a picture of an object from among some other objects has been presented. In addition, the user interface of this method is multilingual. All of the messages are shown in user language so he/she can pass the test even he/she cannot read an English sentence. The structure of the suggested method is very simple and as follows:

At first a bank of the names of some objects is prepared. These names include cars, animals, different persons, flags of countries, etc. The concerned program chooses some of these objects (e.g. 6 objects) randomly. Then the program downloads the pictures about these topics from the Internet. The pictures are searched by Yahoo! image search API [13]. Then all the pictures are merged together and the resulted image is 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. For example the program shows four pictures, car, rabbit, a few kids, and a bird and asks the user to click on the picture of the rabbit. If the user clicks on the picture of the rabbit 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.

Awareness of the shape of the object is very difficult and even impossible due to the great number of the objects in the world. If the program can do the first stage successfully then the recognition of the place of the object is difficult because the program does not know anything about the color, dimensions, and other characteristics of the picture of the object. On the other hand the pictures of the objects have been rotated. As a result the possibility of detection of the place of the concerned picture by the computer is minimized unless the computer randomly chooses a place which is the same as the place of the concerned object but we can minimize this possibility by adding to the number of the pictures.

As mentioned earlier, the user interface of this method is multilingual. At the beginning of the test, a list of supported languages is shown to user. Then the user selects his/her native language. From this point, all of the messages will be shown in user's native language.

Whenever a message must be shown to user, the message is translated to user's native language. The Babel Fish online translator is used to translate the messages [14]. The desired string is sent to the Babel Fish translator and the response is shown to the user.

In current implementation of this method, seven languages are supported: English, Dutch, French, German, Italian, Portuguese and Spanish.

Maybe in the first glance our method seems to be like the method II.A (Implicit CAPTCHA) but these methods are different from each other. Although one implementation of Implicit CAPTCHA method under the name "Single-Click CAPTCHAs" is similar to our method.

That method is based on showing of a series of words, which have various pictorial effects, and in random places on the screen and asking for click on a particular word by user. For example in figure 3 the user is asked to click on the word SUBMIT.

In spite of many similarities between the aforesaid method and ours, that method is based on the OCR in which the user must recognize the words and then detects their places.

Nowadays some methods have been invented which have been successful in recognition of these words in many cases [2], [3] while in our method we deal with showing the picture of the objects and the OCR methods have not been used at all. Furthermore the recognition of a picture, as explained earlier, is more difficult for a computer.



Fig. 3. Implicit CAPTCHA method [4]

Another method which has been presented in "Single-Click CAPTCHAs" section is to show a picture and asking the user to click on a particular place on it. For example a scene of a mountain is shown and the user is asked to click on the mountain top. In this method the concerned points must be determined manually. As a result this method is very time consuming while in our method the pictures are chosen randomly and we don't have to determine the points manually. Furthermore in our method some pictures of different objects are shown and we must choose the picture of the concerned object. On the other hand in our method the pictures are rotated in order to complicate the recognition of the pictures for computer more.

In general it can be said that our suggested method is the improved version of the Implicit CAPTCHA method in which we have effected many changes and combined it with the PIX method[8].

Our method has been implemented by PHP programming language. In this method at first a bank of the names of some objects is created. In addition to the name of the objects, a search string is saved for each object to improve the images founded during web search. For example the string "apple fruit" is stored to be used for finding apple images. The program chooses the names of six objects randomly and the pictures of these objects are downloaded from the Internet. These objects' pictures are searched by Yahoo! image search. The Yahoo! thumbnail cached versions of original pictures are used to decrease the downloading time.

Then the program places the images in a single image in random order. In the merging phase, none of the images will overlap with each other. All of these works are done via image processing functions which are available in standard PHP. Finally the program chooses one of the shown objects randomly and asks the user to click on the picture of that object (Figure 4).

If the user chooses the right object and clicks on it, a message which reads as the successful response of the user is shown and he/she will be allowed to carry out the concerned operations.

A sample implementation of our project is available at <http://www.hip.ir/MultilingualCAPTCHA/>.

Figure 5 shows the page which user sees when he/she passed the test and his/her language is Spanish.

Figure 6 shows the page which user sees when he/she doesn't click on the correct object and his/her language is French.

IV. ADVANTAGES

1. To pass this test, the user does not need any knowledge of English language because all of the messages are shown in his/her native language.
2. Unlike the other CAPTCHA methods which are OCR-based, this method only needs to click on a picture. Sometimes even human users have difficulty in recognizing the displayed words but such a problem does not exist in this method. Also our CAPTCHA method is easy to use and takes little time to pass while not bothering the users.
3. This method is implemented entirely in PHP language. In addition, because PHP is a well known open source and platform independent language and most of the web hosting companies support it, so this method can be easily integrated to available websites with low cost.
4. No keyboard is needed in this method. Therefore, it can be run on devices without a keyboard or devices in which it is difficult to use their keyboards, such as PDA's or mobile phones.
5. It can be used by all ages, even children.

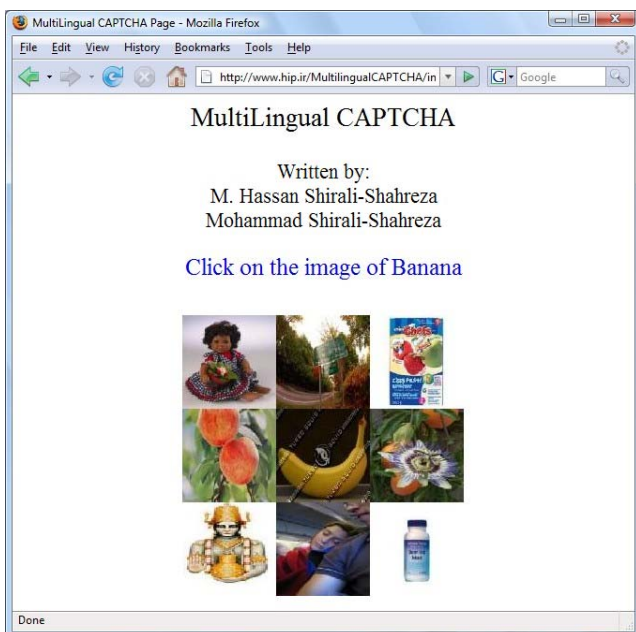


Fig. 4. The Multilingual CAPTCHA method

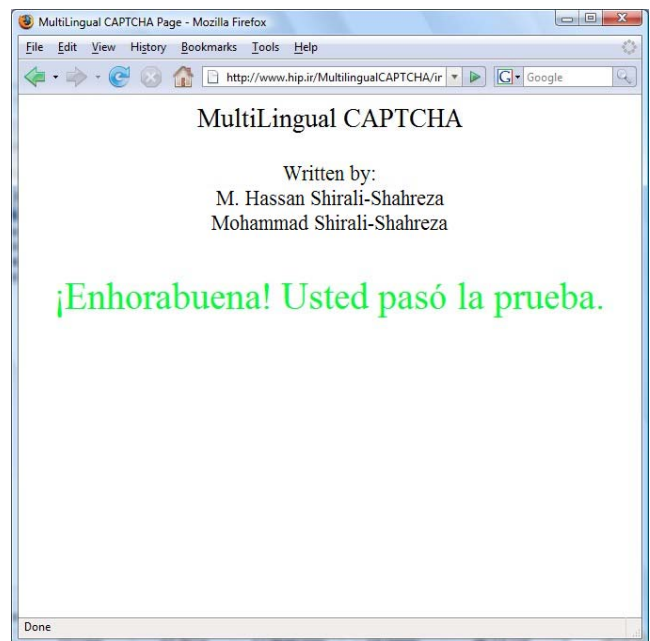


Fig. 5. Test passed message in Spanish



Fig. 6. Test failed message in French

6. By using Internet for downloading pictures, the costs are reduced, because there is no need to create a bank of pictures. Making a local database for images needs more spaces for saving pictures and also needs a lot of time and other resources for creating it.
7. It does not require any processing and can be used on small devices and devices with limited resources such as mobile phones.
8. We can change the difficulty of this method according to its situation. As a result we can make it more difficult by increasing the objects present in the screen, decreasing the distance between the objects and using less familiar objects in a page. This way the recognition and finding of the objects will be more complicated and more difficult for a computer. These operations will make the work more difficult even for a human user.

V. CONCLUSION

This paper presents a method for telling the human users and computer programs apart by showing some pictures simultaneously and asking the user to choose the picture of certain object. Since this method needs two stages (recognition of the object and finding the object) and each of these two operations cannot be done by computer appropriately, this method can resist the computerized attacks efficiently.

On the other hand the users can work with this method easily because all of messages are shown in their native language.

By selecting the pictures from Internet, the cost of the project is reduced and also the varieties of pictures are increased.

We can also implement this method on other devices such as mobile phone, PDA (Personal Digital Assistant), and the devices which have touch screens.

Besides its advantage of covering every age group we can also recommend it for disabled people.

We can make this method specialized for specific websites for example we can use the specialized medical pictures (such as anatomical pictures) for medical websites in order that a user can recognize these pictures.

As a whole our method is versatile and flexible and can be adjusted and modified for various purposes.

REFERENCES

- [1] A. Turing, "Computing Machinery and Intelligence," *Mind*, vol. 59 (236), pp. 433-460, 1950.
- [2] G. Mori, and J. Malik, "Recognizing Objects in Adversarial Clutter: Breaking a Visual CAPTCHA," *Proceedings of the Conference on Computer Vision and Pattern Recognition*, Madison, 2003, pp. 134-141.
- [3] G. Moy et al, "Distortion Estimation Techniques in Solving Visual CAPTCHAs," *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, 2004, pp. 23-28.
- [4] H. S. Baird and J. L. Bentley, "Implicit CAPTCHAs," *Proceedings of the SPIE/IS&T Conference on Document Recognition and Retrieval XII (DR&R2005)*, San Jose, 2005, pp. 191-196.
- [5] T.Y Chan., "Using a Text-to-Speech Synthesizer to Generate a Reverse Turing Test," *Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence*, 2003, pp. 226-232.
- [6] M. Shirali-Shahreza and S. Shirali-Shahreza, "Drawing CAPTCHA," *Proceedings of the 28th International Conference Information Technology Interfaces (ITI 2006)*, Cavtat, Dubrovnik, Croatia, June 19-22, 2006, pp. 475-480.
- [7] H.S. Baird and K. Popat, "Human Interactive Proofs and Document Image Analysis," *Proceedings of the 5th IAPR International Workshop on Document Analysis Systems*, Princeton, LNCS 2423, 2002, pp. 507-518.
- [8] M. Chew and H. S. Baird, "BaffleText: a Human Interactive Proof," *Proceedings of the 10th SPIE/IS&T Document Recognition and Retrieval Conference (DRR2003)*, Santa Clara, CA, 2003, pp. 305-316.
- [9] Paypal, PayPal registration, <https://www.paypal.com>
- [10] H. S. Baird and T. Riopka, "ScatterType: a Reading CAPTCHA Resistant to Segmentation Attack," *Proceedings of the IS&T/SPIE Document Recognition & Retrieval XII Conference*, CA, 2005, pp. 197-207.
- [11] A.L Coates et al, "Pessimial Print: A Reverse Turing Test," *Proceedings of the 6th International Conference on Document Analysis and Recognition*, WA, USA, 2001, pp. 1154-1158.
- [12] M.H. Shirali-Shahreza and M. Shirali-Shahreza, "Persian/Arabic BaffleText CAPTCHA," *Journal of Universal Computer Science (J.UCS)*, vol. 12, no. 12, pp. 1783-1796, December 2006.
- [13] Yahoo!, Yahoo Web Search API, <http://developer.yahoo.com>
- [14] Altavista, Babel Fish Translation, <http://babelfish.altavista.com>