ISSN:0975-9646

Siddharth Syal / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (2) , 2016, 517-518

Alternatives to CAPTCHA

Siddharth Syal

Computer Science and Engineering
SRM University, Chennai, India

***Abstract* – In order to reduce the problem of bots filling up various forms over the internet and sending unnecessary request to the servers we generally use CAPTCHA. CAPTCHA stands for Completely Automated Public Turing test. It is used to stop the bots from automatically filling the form or filling up the servers with unnecessary request. Originally, CAPTCHA used images and asked the humans to write the text from the image to a text box and then the computer would believe that a human has filled up the form, but due to development of OCR technologies, bots can also read from the images and fill up the forms and send the requests. Various alternatives like measuring the time used by bots and humans to fill a form, Task puzzles and Honeypots can help, in a better way, to stop automated form filling and automated requests.**

Keywords - : CAPTCHA, OCR, bots, alternatives, turning test.

I. WHAT IS CAPTCHA?

CAPTCHA is the short name for Completely Automated Turning Test. This mechanism is used by website in order to avoid the automated form filling and automated HTTP request to the servers. The conventional form of CAPTCHA (fig 1.1) has an image that is distorted in such a way that humans can understand what is the content of the image and then the human is asked to write the content of the image in a text box on the web page and then the computer confirms that it's human that has filled the form or sent a request. But due to recent development of OCR technologies, computer programs can be designed in such a way that they can read the information that is present in the images, due to which they can fill up the text box easily and this helps to bypass the protocols that are used to stop bots and spam content.

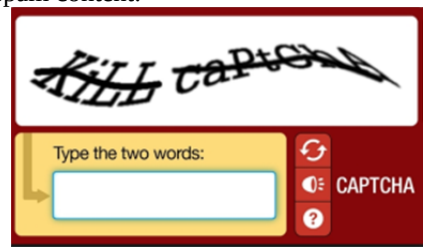

Fig 1.

For users who are void of clear vision capabilities, an audio message is played and the content of that message is then reproduced in the text box which is provided in the webpage.

II. PROBLEMS WITH CAPTCHA

There are various cons of the conventional CAPTCHA that is used. Firstly, due to the development of the OCR technologies, computers can easily read scanned data and can reproduce the same data by itself. So, when the image is produced by the CAPTCHA, the bots can be programmed to use the OCR technology and fill out the text box with appropriate content and can easily bypass the spam prevention mechanism. Secondly, the conventional CAPTCHA can sometimes be annoying as the letters could be distorted in such a way that even humans cannot read the text properly.

III. ALTERNATIVES TO CAPTCHA

Due to the development of various new technologies, the conventional CAPTCHA is not able to provide security against the internet bots so various alternatives have been suggested. Following are the alternatives to CAPTCHA - :

- **Timing Trick –** We can measure the average time taken by the human and a bot to fill up a form and then use those results to confirm if a human is trying to interact with the computer.
- **The Honeypot –** Various fields can be introduced in a legit form and those field are designed using JavaScript or CSS, so that only bots can see them and humans cannot see them, so if those fields are filled, then it can be confirmed that a bot is trying to send request to the server or trying to fill a form.
- **Puzzle Games –** We can introduce simple tasks in form of puzzles that are interactive for the humans and require little intelligence to fulfill the needs of the task but it will be difficult for a bot to finish the task.

IV. TIMING TRICK

This is one of the easiest ways to stop spam bots. In this method we generate a function that randomly measures the time taken by a bot and a human to fill up a form of desired length and the average of those values are noted. Then these values are used whenever a real form is accessed over the internet. After the real form is filled, the time taken to fill the form is noted and measured with the values that were calculated by the function. If the values are nearby the values measured by the function, an appropriate action is taken. Generally, on an average, a human will take a minute or two to fill up a form of average length, but a bot will take few seconds to fill a form of any length.

Although this way is the easiest to implement but it has its own short comings. Various browsers like Google Chrome and Safari have a feature for auto-filling the form, but these features require the permissions of the users in order to fill the form. So if the website uses the timing trick feature, then in that case the website will treat the form filled using auto-fill feature same as the form filled by bots or some spam generation mechanism and a legit user will be not allowed to access the website or send any request to the server.

V. THE HONEYPOT

Honeypot is a method in which the spam bots are tempted towards a particular field in a form while the same field remains invisible to a normal user. In this method, a field is created on the website and the field is kept hidden from the normal users with the help of CSS and JavaScript, but the bots can easily identify the field. Once the bot identifies the field, it will try to add some data to the field and submit the form or create a server request. In the case of honeypot, once the form or a request is submitted, the hidden field is checked to see if that field is filled or not. If the field is filled,

then the computer knows that it was a bot that had filled the form and not a human being and then the access is denied.

```
<div id="honeypotsome-div">
If you see this, leave this form field blank
and invest in CSS support.
<input type="text" name="body" value="" />
</div>
```

Fig 2.

For example (fig 2), in order to make a honeypot in a website, we can make a form field named 'body' assuming that the form field which is visible to users is named as 'real body'. This field named 'body' will remain invisible to users while it will be visible to the bots.

```
if(!String.IsNullOrEmpty(Request.Form["body"]))
  IgnoreComment();
```

Fig 3.

Later, when the form or any request is submitted, a code (Fig 3) can be made in order to check if the field is empty or contains any data.

While generating the form field for honeypot, form field names such as hide, hidden etc should be avoided, as the programmers can make bots that will first scan the form field name and if it has names like hide, then the bots may simply ignore it. So form field names should be chosen in such a way that the field looks like a part of actual form.

VI. PUZZLE GAME

It is generally seen that introducing CAPTCHA to any website, increases the number of drop offs from the website. The major reason behind the drop offs is that people find it cumbersome to fill in the text given in the CAPTCHA and even though they might fill in the text properly, the CAPTCHA may produce false negatives. In order to make the process of human verification more interactive we can use puzzle games so that the human constantly interacts with the computer in order to verify if it's a bot or a human interacting with the system.

The major difference between a human and a well programmed bot is the level of intelligence. Using puzzle game as a way of verifying exploits the human intelligence in such a way that not all bots can solve the puzzle easily.

Fig 4.

We can make use of certain basic tasks like putting the food items in the plate (Fig 4). The tasks like these require minimal human intelligence and various recent technologies like OCR cannot be used in such cases. The developer of these types of CAPTCHA's should always randomly select the amount of elements that are to be dragged from one place to another. Moreover, the session ID of each of the iteration of CAPTCHA should be changed so that a programmer cannot develop a bot to solve the same CAPTCHA.

The major disadvantage of using such a method for human verification is that the programmer may need a large image database for generating the puzzle games because if a small database is used for generating the puzzle game, it can be the case the bot programmer can use various techniques to find all the images in the database and then generate a bot in order to solve those CAPTCHA s.

VII. CONCLUSION

The conventional CAPTCHA is not able to provide a good level of protection against spam bots, so various new mechanisms are suggested. The new mechanisms can be used with various combinations like puzzle game with timing trick, so that the level of protection is enhanced and the process is less cumbersome to the normal user.

REFERENCES

[1] https://www.capy.me/products/puzzle_CAPTCHA /
[2] http://www.sitepoint.com/better-captcha/
[3] http://www.scientificamerican.com/article/pogue-8-alternatives-to-hated-captcha/
[4] http://www.dailydot.com/technology/end-captcha-security-top-five-alternatives/
[5]https://www.w3.org/WAI/GL/wiki/Captcha_Alternatives_and_thoughts