

Shevchuk Bogdan

*Student of the Faculty of Informatics and
Computer Science of the
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"*

WHAT YOU SHOULD KNOW ABOUT CAPTCHAS

***Summary.** Technology nowadays is a part of a life almost every human on earth, creating problems of different sorts of cause. While surfing an internet not many are concerned about their privacy online. To regulate all the information getting into the world-web was created whole package of security measures and instruments. One of them was CAPTCHA.*

***Key words:** CAPTCHA, Safety, Protection, Security, AI, Experimentation.*

What is CAPTCHA?

CAPTCHA is acronym for «completely automated public Turing test to tell computers and humans apart» which implies that this is the test to tell if the user isn't a malicious program. Trying to prevent these bots entering the system and making their service more secured in process was invented CAPTCHA, simple test that required to retype slightly distorted symbols from the picture when user is trying to log in [1].

Why do we need it?

Nowadays enormous amount of people are using internet in their routine life without even noticing how fragile their privacy is. There are a lot of dangerous people out there who have a job trying to get any dip of information about their prey. Big companies like Google, Facebook, PayPal, etc. are very

concerned about privacy of their users so they have taken some measures to protect their data. By forcing every user to pass this test companies are lowering cases of losing private information into wrong hands.

There are many types of CAPTCHA used by different sources, but we will investigate only a few.

- reCAPTCHA

Technology acquired by Google in 2009 that is evolving every time when someone uses it. It was developed to work in tandem with their own AI. The idea behind it is simple. Google has started a new program to digitalize every single scanned file like books, journals, words on photos etc. When AI could not understand what word was scanned, it just crops it and sends to a user. reCAPTCHA is forcing user to give the right answer without knowing, that he helped AI to recognize text [2].

- No CAPTCHA reCAPTCHA

A new branch in Google's technology of reCAPTCHA. It is the most up-to-date software that this company could offer. To make it easier and quicker to pass this test they have developed new system, which tracks your behavior before entering the system. If system "thinks" that you are suspicious it forces you to pass the test. Test is slightly different from the previous one. By using same tricks, Google is trying to make people to teach AI understanding of what was depicted on pictures and photos. By using this method, they can categorize every digital image for more convenient uses. To simply add this type of CAPTCHA to your system, Google allows to simply add these strings to their page.

```
<html>
<head>
<title>reCAPTCHA demo: Simple page</title>
<script src="https://www.google.com/recaptcha/api.js" async
defer></script>
```

```
</head>  
<body>  
<form action="?" method="POST">  
<div class="g-recaptcha" data-sitekey="your_site_key"></div>  
<br/>  
<input type="submit" value="Submit">  
</form>  
</body>  
</html>
```

The result will look similarly to this Fig.1:

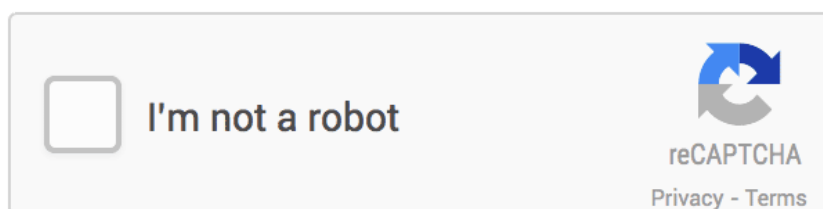


Fig. 1. The NoCAPTCHA reCAPTCHA

- 3D CAPTCHA

This example has the same principle as text based CAPTCHA, but instead of making a distorted image of symbols it generates a 3D version of them. By using this technique text will be easier to understand to user while being hard enough to be passed by a machine. This type of test is more flexible for distortion and editing because of dimensions there are. This type of CAPTCHA is not popular on the internet because of the process it is required to go through.

- Math solving CAPTCHA

A simple test requires you to solve some simple math exercise to pass through. Although it is an easy task for everyone, you can imagine how easily it bypassed by non-human users. It will look similarly like this Fig. 2:

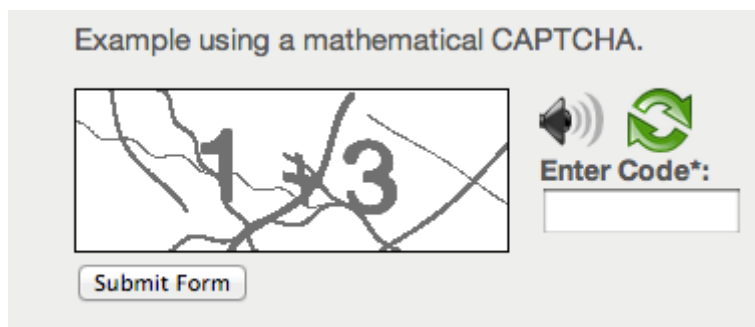


Fig. 2. The Math solving CAPTCHA

- 3D CAPTCHA based on spatial perspective and human imagination

A new type of CAPTCHA technology based on spatial perspective and human imagination[3]. The idea behind this is to find the correct angle of special 3D object to find the solution. Using original 2D image program is generating a special formed 3D object. The method is described better on Fig. 3.

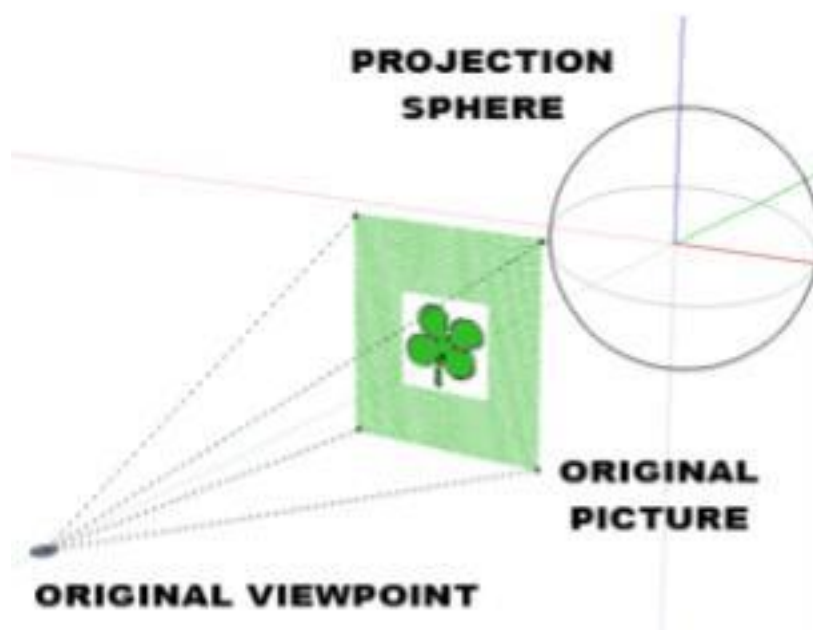


Fig. 3. Generating the 3D object

By dividing the picture into several parts and moving them into another dimension it creates illusion of 2D object only from one single angle. By tasking user to rotate volume object this program doesn't create big problems in

understanding how to use it. The task is to find the same right observation point. For comparison computers could not understand which point is right because it is creating illusion only for human eye. The result presented on Fig. 4.



Fig. 4. 3D CAPTCHA testing

How secure it is to use CAPTCHA?

Since the internet went online, the question about security was present. Various companies tried to protect their users from spammers and bots to prevent info leaking. While the technology to protect information was evolving, internet scammers were not just standing around. Every technology has its weakness and someone will eventually find it.

While developing their measures of security companies has met another problem from another side, the user himself. People with some sort of dysfunction could not see or just understand what is on the screen. For these

people was added voiced CAPTCHA that allowed a person with bad sight to pass the test.

Eventually these tests was bypassed with already evolved bots or just services that solved these tests for you. With time these tests became more and more hard to understand by real human so services, that allowed solving them for you became more popular.

Trying to be more secured big companies invested a lot of money to investigate this problem. They created various types of tests with AI that was developed by them in a process [4]. As a result approximately 80% of people couldn't pass these tests, while AI solved 99% of them.

To create a reliable type of security Google announced their solution. By tracking users behavior before clicking on a button to sign in their AI can distinguish human from a robot. You see, people move their cursors in unideal, wiggly way, which cannot be similar to bot behavior. Their machines also process information about what sites you previously visited and on which parts of page you were looking the most. It was made to calculate whether you are a real human. It is very suspicious for this company to store these types of information. Unfortunately, they do not give a full list of data they use to measure your "humanity". Many people are concerned about what kind of data are they sharing and where it is used.

Behaving like this companies cause popularity of services that allows solving CAPTCHAs for you. Thousands of people in third world countries are making the living of these services [5]. Every second they receive simple test to solve for someone who has no desire to solve these by any cause but have money to make others to do it for them. For comparison, one million solved CAPTCHAs are worth one thousand dollars. For some people it is not worth it, but for someone it is the way to make a living.

In trying to gain access to specific data, criminals are developing their software to bypass all these security. Every year every piece of software are in

danger of being hacked or stolen. While some are using living force to gain access to services, others tempted to develop own hardware and use it to sell its capabilities [6]. Every time when there is new tech on the horizon, there are someone who already trying to hack it. For every CAPTCHA there is exists bypassing tool, allowing hackers to momentarily enter any system without being noticed. This can be used for various reasons like:

- Buying something in big quantities in short period of time to resell later for more money;
 - DDOS attacks;
 - Creating accounts for mass spam;
 - Phishing private information of users;
 - Forcing virus links as trusted;
 - Advertisement;
- Etc.

There are no ultimate security measures to divide the consumer from the abuser or a criminal. Each day engineers from all over the world is trying to evolve their software, adapt it in any way, so it would be harder to be manipulated, but easy to use. Each time you check the mark, proving you are a human, you need to remember, that on this checkmark worked thousands of people to make it as accessible and secure as possible.

References

1. CAPTCHA – Retrieved from <https://en.wikipedia.org/wiki/CAPTCHA>
2. ReCAPTCHA – Retrieved from <https://en.wikipedia.org/wiki/ReCAPTCHA>
3. 3D CAPTCHA – Retrieved from http://www.3dcaptcha.net/documents/3D_captcha.pdf
4. What is reCAPTCHA? - Retrieved from <https://developers.google.com/recaptcha/intro>

5. WHAT IS CAPTCHA?– Retrieved from <http://chris.pirillo.com/2011/06/10/what-is-captcha/>
6. CAPTCHA - Retrieved from <https://captcha.org/>