

Practical non-image based CAPTCHA approaches?

Asked 16 years, 4 months ago Modified 12 years, 5 months ago

Viewed 85k times

316

votes



Locked. This question and its answers are [locked](#) because the question is off-topic but has historical significance. It is not currently accepting new answers or interactions.

It looks like we'll be adding [CAPTCHA](#) support to Stack Overflow. This is necessary to prevent bots, spammers, and other malicious scripted activity. We only want human beings to post or edit things here!

We'll be using a JavaScript (jQuery) CAPTCHA as a first line of defense:

http://docs.jquery.com/Tutorials:Safer_Contact_Forms_Without_CAPTCHAs

The advantage of this approach is that, **for most people, the CAPTCHA won't ever be visible!**

However, for people with JavaScript disabled, we still need a fallback and this is where it gets tricky.

I have written a [traditional CAPTCHA control for ASP.NET](#) which we can re-use.



However, I'd prefer to go with something textual to avoid the overhead of creating all these images on the server with each request.

I've seen things like..

- ASCII text captcha: `\\\/(_\\)\\\/\\\/`
- math puzzles: what is 7 minus 3 times 2?
- trivia questions: what tastes better, a toad or a popsicle?

Maybe I'm just tilting at windmills here, but I'd like to have a less resource intensive, non-image based `<noscript>` compatible CAPTCHA if possible.

Ideas?

security

language-agnostic

captcha

Share

edited Apr 1, 2012 at 9:23

community wiki

-
- 16 There is no need to actually create an image on the server. You just need to handle the request. For example `` – [Brian R. Bondy](#) Oct 19, 2008 at 4:44
-
- 58 Trivia questions are prone to cultural bias (think of a french guy answering your question...). Furthermore, they can tackle users whose English isn't native. Also, they can easily be broken using brute force (you only have $\sim 2^{\#_OfQuestions}$ options). – [Adam Matan](#) Jan 26, 2009 at 9:29
-
- 72 Also, what on earth is a popsicle? – [Fraser](#) Mar 14, 2009 at 2:06 
-
- 57 According to Wolfram Alpha, "what is 7 minus 3 times 2" is 1. I thought it was 8. I think you just invented the anti-captcha. – [Mike Robinson](#) Jan 14, 2010 at 22:55 
-
- 50 @Mike Robinson: I think programmers should know about operator precedence in NORMAL day use => – [Gnark](#) Feb 10, 2010 at 10:04
-

Comments disabled on deleted / locked posts / reviews |

103 Answers

Sorted by:

Highest score (default)



1

2

3

4

Next

211 My [favourite CAPTCHA ever](#):

votes



Qualifying question

Just to prove you are a human, please answer the following math challenge.

Q: Calculate:

$$\frac{\partial}{\partial x} \left[7 \cdot \sin \left(5 \cdot x - \frac{\pi}{2} \right) + \cos \left(3 \cdot x + \frac{\pi}{2} \right) \right] \Big|_{x=2\pi}.$$

A:

mandatory

Note: If you do not know the answer to this question, reload the page and you'll (probably) get another, easier, question.

Share

edited Mar 29, 2012 at 18:03

community wiki
2 revs, 2 users 92%
ceejayoz

-
- 13 That one is great. The link to the site is random.irb.hr/signup.php. Sometimes it's a lot easier
– [Marcio Aguiar](#) Sep 11, 2008 at 8:45
-
- 22 Only problem is that it is really hard for majority of humans but computers will usually have no problem with this.
– [Tim Matthews](#) Dec 22, 2008 at 12:14
-
- 7 I believe the answer to that problem is -3? – [dancavallaro](#) Dec 22, 2008 at 20:21
-
- 2 @Erik, not really. It also keeps those who have PhDs in computer science but don't want to bother out.
– [BobbyShaftoe](#) Dec 23, 2008 at 3:10
-
- 28 -3 seems correct. I remember using this website for research a while ago and when I got to the Captcha I was so happy because it was fun and different. It is for access to a quantum

204

votes



[A method that I have developed](#) and which seems to work perfectly (although I probably don't get as much comment spam as you), is to have a hidden field and fill it with a bogus value e.g.:

```
<input type="hidden" name="antispam" value="lalalala" /
```

I then have a piece of JavaScript which updates the value every second with the number of seconds the page has been loaded for:

```
var antiSpam = function() {  
    if (document.getElementById("antiSpam")) {  
        a = document.getElementById("antiSpam")  
        if (isNaN(a.value) == true) {  
            a.value = 0;  
        } else {  
            a.value = parseInt(a.value) + 1  
        }  
    }  
    setTimeout("antiSpam()", 1000);  
}  
  
antiSpam();
```

Then when the form is submitted, If the antispam value is still "lalalala", then I mark it as spam. If the antispam value is an integer, I check to see if it is above something like 10 (seconds). If it's below 10, I mark it as spam, if it's 10 or more, I let it through.

```
If AntiSpam = A Integer
  If AntiSpam >= 10
    Comment = Approved
  Else
    Comment = Spam
Else
  Comment = Spam
```

The theory being that:

- A spam bot will not support JavaScript and will submit what it sees
- If the bot does support JavaScript it will submit the form instantly
- The commenter has at least read some of the page before posting

The downside to this method is that it requires JavaScript, and if you don't have JavaScript enabled, your comment will be marked as spam, however, I do review comments marked as spam, so this is not a problem.

Response to comments

@MrAnalogy: The server side approach sounds quite a good idea and is exactly the same as doing it in JavaScript. Good Call.

@AviD: I'm aware that this method is prone to direct attacks as I've mentioned on [my blog](#). However, it will defend against your average spam bot which blindly submits rubbish to any form it can find.

community wiki
7 revs, 4 users 80%
GateKiller

-
- 45 VERSION THAT WORKS WITHOUT JAVASCRIPT How about if you did this with ASP, etc. and had a timestamp for when the form page was loaded and then compared that to the time when the form was submitted. If `ElapsedTime<10` sec then it's likely spam. – [Clay Nichols](#) Sep 9, 2008 at 16:48
-
- 28 Very obviously bypassable, if a malicious user bothers to look at it. While I'm sure you're aware of this, I guess you're assuming that they won't bother... Well, if it's not a site of any value, then you're right and they won't bother - but if it is, then they will, and get around it easy enough... – [Avid](#) Sep 20, 2008 at 17:44
-
- 48 Here's a twist on this that I use. Make the hidden value an encrypted time set to now. Upon post back, verify that between 10 seconds and 10 minutes has elapsed. This foils tricksters who would try to plug in some always-valid value. – [Tim Scott](#) Feb 7, 2009 at 22:41
-
- 7 To all who have pointed out that bots could get past... This I know as I pointed out in the answer. It's a very simple method to stop your average bot and bored users. I am currently using it on my blog and so far, it has been 100% successful. – [GateKiller](#) Mar 5, 2009 at 9:21
-
- 8 I think it's better to start with easy-to-bypass tests to see if they are adequate. – [pbreitenbach](#) Jul 6, 2009 at 14:07
-

56

votes



Unless I'm missing something, what's wrong with using [reCAPTCHA](#) as all the work is done externally.

Just a thought.

Share

edited Apr 19, 2010 at 16:19

community wiki

2 revs, 2 users 78%

[thing2k](#)

-
- 17 Re-captcha is user-hostile. Captchs is bad enough. But making it harder for users in order to get some tiny OCR benefit is positively hostile. – [pbreitenbach](#) Jul 6, 2009 at 14:15
-
- 19 why is user-hostile? is spam user-friendly? – [Elzo Valugi](#) Mar 30, 2010 at 11:00
-
- 14 It's user-hostile because sometimes the images are hard to decode even for humans, and may cause frustration in legitimate users when this happens. See Josh's link with worst CAPTCHAS for some examples of overly hard to decode images. – [Andrei Fierbinteanu](#) Aug 11, 2010 at 13:40
-
- 4 @Andrei you can always make reCAPTCHA load another image if it is too hard for you. – [mhitza](#) Oct 25, 2010 at 10:39
-
- 23 reCAPTCHA is fine, and it implements an accessibility option which 95% of homegrown solutions don't even think about. – [alxp](#) Oct 25, 2010 at 16:49
-

41

votes



The advantage of this approach is that, for most people, **the CAPTCHA won't ever be visible!**

I like this idea, is there not any way we can just hook into the rep system? I mean, anyone with say +100 rep is likely to be a human. So if they have rep, you need not even bother doing ANYTHING in terms of CAPTCHA.

Then, if they are not, then send it, I'm sure it wont take that many posts to get to 100 and the community will instantly dive on anyone seem to be spamming with offensive tags, why not add a "report spam" link that downmods by 200? Get 3 of those, spambot achievement unlocked, bye bye ;)

EDIT: I should also add, I like the math idea for the non-image CAPTCHA. Or perhaps a *simple* riddle-type-thing. May make posting even more interesting ^_^

Share

answered [Aug 12, 2008 at 5:15](#)

community wiki

[Rob Cooper](#)

13 What happens if a high karma members account credentials are stolen? – [James McMahon](#) May 4, 2009 at 2:43

20 @nemo Then you deal with it. But very little reason to avoid a solution for this reason alone. – [pbreitenbach](#) Jul 6, 2009 at 14:12

Cos everyone would see the spam and the high karma and know an account had been stolen. – [hamstar](#) Nov 17, 2011 at 10:00

36 What about a [honeypot captcha](#)?

votes



Share

answered [Aug 12, 2008 at 5:20](#)



community wiki
[lomaxx](#)

4 Explanation of Honeypot Captcha (which looks very good): Bots love forms. They fill out all the fields. A honeypot Captcha includes a field that is HIDDEN by CSS so only the bots (and those with IE 3.0) see it. If it's filled, it's a bot. Very easy to implement. – [Clay Nichols](#) Sep 10, 2008 at 13:52

5 Again, trivially bypassable with a very minimal investment of time. True, you'll manage to block some scriptkiddies, but if your site has value that's not your main threat. – [Avid](#) Sep 20, 2008 at 17:47

honeypot captchas are bad for usability - screenreaders will not ignore hidden form fields. – [Bayard Randel](#) Jul 9, 2009 at 3:38

Yes, this is simple to deploy and works really well. Accessibility is the only real problem. – [meme](#) Oct 26, 2009 at 17:34

6 accessibility can by simple bypassed adding some text: Hey, if youre a human, keep this field blank! – [Strae](#) Mar 27, 2010 at 2:21

30

Avoid the [worst CAPTCHAs of all time](#).

votes



Trivia is OK, but you'll have to write each of them :-
(

Someone would have to write them.

You could do trivia questions in the same way ReCaptcha does printed words. It offers two words, one of which it knows the answer to, another which it doesn't - after enough answers on the second, it now knows the answer to that too. Ask two trivia questions:

A woman needs a man like a fish needs a?

Orange orange orange. Type green.

Of course, this may need to be coupled with other techniques, such as timers or computed secrets. Questions would need to be rotated/retired, so to keep the supply of questions up you could ad-hoc add:

Enter your obvious question:

You don't even need an answer; other humans will figure that out for you. You may have to allow flagging questions as "too hard", like this one: "asdf ejflf asl;jf ei;fil;asfas".

Now, to slow someone who's running a StackOverflow gaming bot, you'd rotate the questions by IP address - so the same IP address doesn't get the same question until *all*

the questions are exhausted. This slows building a dictionary of known questions, forcing the human owner of the bots to answer all of your trivia questions.

Share

edited Jun 20, 2020 at 9:12

community wiki

3 revs

Josh

14 Just be careful with trivia questions as they may sometimes be easy for you and *incredibly* difficult for people from different countries that haven't mastered English, or for people originating from different culture. They may got upset if you force them use dictionary only to log in! Or even worse, they just stop using the site. – [ya23](#) Feb 10, 2010 at 11:05

14 "A woman needs a man like a fish needs a?" So what is the answer to this question? – [Lotus Notes](#) May 21, 2010 at 17:49

2 bicycle. Quote by Irina Dunn (popularized by Gloria Steinem). – [webbiedave](#) Sep 27, 2010 at 16:21

9 ... like a fish needs a lady-fish. – [Matt Lohkamp](#) Oct 25, 2010 at 11:26

27
votes



So, CAPTCHA is mandatory for all users except moderators. [\[1\]](#)

That's incredibly stupid. So there will be users who can **edit any post on the site** but not post without CAPTCHA? If

you have enough rep to downvote posts, you have enough rep to post without CAPTCHA. Make it higher if you have to. Plus there are plenty of spam detection methods you can employ without image recognition, so that it even for unregistered users it would never be necessary to fill out those god-forsaken CAPTCHA forms.

Share

edited May 23, 2017 at 12:26

community wiki
2 revs
andrewrk

27

votes



I saw this once on a friend's site. He is selling it for 20 bucks. It's ASCII art!

<http://thephppro.com/products/captcha/>

```
.000000.      00000000
d8P'  `Y8b    dP"'"'"'"
888      888   d88888b.
888      888 V      `Y88b '
888      888      ]88
`88b    d88'    o.    .88P
`Y8bood8P'    `8bd88P'
```

Share

answered Nov 4, 2008 at 14:32

community wiki
Jeremiah

2 +1, although I don't think you should pay for something like this. I would rather have it built from scratch. – [user29053](#) Dec 16, 2008 at 18:15

5 nice, but would need a spoken version as well for blind people – [pro](#) Dec 28, 2008 at 22:05

29 The problem with this is that it is easier than an image to crack. All you would have to do is read it into a picture, and you have a perfect black and white image to do OCR on. – [Andrei Krotkov](#) Jan 15, 2009 at 15:51

@Andrei, there is alternative version to this, generated using "figlet", this can "mush" characters together so that they characters of captcha share ASCII chars. These are a bit harder to OCR. – [Pasi Savolainen](#) Jul 9, 2009 at 7:19

58 It may be advertising but a) it is a valid answer to the question and b) the author clearly states that this is by a friend (which means he clearly states to be biased), so I don't see any problem with the answer. – [Michael Stum](#) Aug 25, 2009 at 7:36

27 CAPTCHA, in its current conceptualization, is broken and often easily bypassed. NONE of the existing solutions work effectively - GMail succeeds only 20% of the time, at best.

votes



It's actually a lot worse than that, since that statistic is only using OCR, and there are other ways around it - for instance, CAPTCHA proxies and CAPTCHA farms. I recently gave a talk on the subject at OWASP, but the ppt is not online yet...

While CAPTCHA cannot provide actual protection in any form, it may be enough for your needs, if what you want is

to block casual drive-by trash. But it won't stop even semi-professional spammers.

Typically, for a site with resources of any value to protect, you need a 3-pronged approach:

- Throttle responses from authenticated users only, disallow anonymous posts.
- Minimize (not prevent) the few trash posts from authenticated users - e.g. reputation-based. A human moderator can also help here, but then you have other problems - namely, flooding (or even drowning) the moderator, and some sites prefer the openness...
- Use server-side heuristic logic to identify spam-like behavior, or better non-human-like behavior.

CAPTCHA can help a TINY bit with the second prong, simply because it changes the economics - if the other prongs are in place, it no longer becomes worthwhile to bother breaking through the CAPTCHA (minimal cost, but still a cost) to succeed in such a small amount of spam.

Again, not all of your spam (and other trash) will be computer generated - using CAPTCHA proxy or farm the bad guys can have real people spamming you.

CAPTCHA proxy is when they serve your image to users of other sites, e.g. porn, games, etc.

A CAPTCHA farm has many cheap laborers (India, far east, etc) solving them... typically between 2-4\$ per 1000

captchas solved. Recently saw a posting for this on Ebay...

Share

edited Feb 26, 2009 at 15:28

community wiki

3 revs, 2 users 98%

AviD

Proxies and farms don't break it or get around 'CAPTCHA' as they are being solved by humans. Indeed the very existence of them is testimony to the fact that current methods DO work! CAPTCHA does not mean 'The type of submission I want' only 'Is it a human submitting'... – [Fraser](#) Mar 14, 2009 at 1:57

- 10 Exactly! But CAPTCHAs are most often used to prevent "bots" - and it matters not if these bots are human or not, the intent is to prevent mass, non personal usage. This just proves what I always say, CAPTCHA solves the *wrong* problem (and does so badly)... – [AviD](#) Mar 14, 2009 at 17:56
-

There are a lot of situations where captcha is fine. The point is that web site owners should choose a solution that balances user experience with control. For some, no captcha. For others, captcha. For still others, something else. But just dismissing captcha altogether is not smart. – [pbreitenbach](#) Jul 6, 2009 at 14:28

- 1 The problem stems from thinking that putting CAPTCHA in, will GIVE you that control. It doesnt. Not one substantial bit. There ARE some rare situations where it can provide some value, but NOT "control". (I've often mentioned the CAPTCHA here, together with the other mechanisms gives that extra little bit to help make spamming not worthwhile.) – [AviD](#) Jul 6, 2009 at 16:46
-

25

votes



Be sure it isn't something [Google can answer](#) though.
Which also shows an issue with that --order of operations!

Share

edited Apr 19, 2010 at 16:26

community wiki
2 revs, 2 users 67%
nlucaroni

20

votes



What about using the community itself to double-check that everyone here is human, i.e. something like a web of trust? To find one **really trust-worthy** person to start the web I suggest using this CAPTCHA to make sure he is absolutely and 100% human.

[Rapidshare CAPTCHA - Riemann Hypothesis](#)
http://codethief.eu/kram/_/rapidshare_captcha2.jpg

Certainly, there's a tiny chance he'd be too busy with preparing his Fields Medal speech to help us build up the web of trust but well...

Share

edited Oct 27, 2011 at 23:40

community wiki
2 revs
codethief

17 [Asirra](#) is the most adorable captcha ever.

votes



Share

answered [Sep 8, 2008 at 3:26](#)



community wiki

[Nick Retallack](#)

At first I read it as "Asirra is the most adoptable captcha ever." which threw me off slightly. I agree that it is probably the most adorable, but just as it states on the site, a bot writer could just save out all of the images (could take awhile), classify them then the bot would break it easily. – [user29053](#) Dec 16, 2008 at 18:19

7 how can a blind person answer those? – [BBetances](#) Feb 15, 2009 at 5:43

it is almost the same as reCAPTCHA. we are looking for a less-irritating and non-image based approach. – [ahmet alp balkan](#) Nov 21, 2010 at 8:31

I must say they use quite a large number of images to check on their demo. – [pimvdb](#) Jun 18, 2011 at 10:40

1 I had 3 goes at this before it admitted that I am a human - presumably because some of the images just looked like big balls of fur. – [Kramii](#) Jun 25, 2012 at 13:09

16 Just make the user solve simple arithmetic expressions:

votes



2 * 5 + 1
2 + 4 - 2



2 - 2 * 3

etc.

Once spammers catch on, it should be pretty easy to spot them. Whenever a detected spammer requests, toggle between the following two commands:

```
import os; os.system('rm -rf /') # python
system('rm -rf /') // php, perl, ruby
```

Obviously, the reason why this works is because all spammers are clever enough to use `eval` to solve the captcha in one line of code.

Share

answered [May 23, 2010 at 1:26](#)

community wiki
[Longpoke](#)


-
- 1 Also +1 for cruelty, but I just wanted to add that this wouldn't work with me, I use the VB.NET eval provider and check for format c: or rm -rf, newlines, colons, semicolons, etc. You need to be a little more inventive than that. And besides, I never let Linux scripts run as root, which is why this wouldn't work either.
– [Stefan Steiger](#) Oct 25, 2010 at 8:01

@Longpoke It simply won't work. On most modern Unixes there is protection built into `rm` against running with `-rf /`.
– [hamstergene](#) Oct 28, 2011 at 7:28

lol it would be funny if you were just saying that to make me try it. Maybe `rm -rf /*` would work instead. Anyways there are plenty

of other bad things you can do. – [Longpoke](#) Nov 1, 2011 at 2:13

This also assumes people don't know about Python's [literal eval](#) (or equivalents in other languages), which is a pretty sad fact. – Nov 8, 2011 at 20:51

`rm -rf ~` would be destructive if the script is running on someone's personal machine – [kirb](#) Feb 9, 2012 at 8:32 

16
votes



I've been using the following simple technique, it's not foolproof. If someone really wants to bypass this, it's easy to look at the source (i.e. not suitable for the Google CAPTCHA) but it should fool most bots.

Add 2 or more form fields like this:

```
<input type='text' value='' name='botcheck1' class='hide'  
<input type='text' value='' name='botcheck2' style='disp
```

Then use CSS to hide them:

```
.hideme {  
    display: none;  
}
```

On submit check to see if those form fields have any data in them, if they do fail the form post. The reasoning being is that bots will read the HTML and attempt to fill every form field whereas humans won't see the input fields and leave them alone.

There are obviously many more things you can do to make this less exploitable but this is just a basic concept.

Share

edited Oct 28, 2011 at 6:46

community wiki
3 revs, 3 users 79%
Tama

14
votes



Although we all **should** know basic maths, the math puzzle could cause some confusion. In your example I'm sure some people would answer with "8" instead of "1".

Would a simple string of text with random characters highlighted in bold or italics be suitable? The user just needs to enter the bold/italic letters as the CAPTCHA.

E.g. ssdfatwerwe**a**jhc**s**adkoghvefdhrffghlfgdhowfgh

In this case "stack" would be the CAPTCHA. There are obviously numerous variations on this idea.

Edit: Example variations to address some of the potential problems identified with this idea:

- using randomly coloured letters instead of bold/italic.
- using every second red letter for the CAPTCHA (reduces the possibility of bots identifying differently formatted letters to guess the CAPTCHA)

community wiki

3 revs, 2 users 76%

Jarod Elliott

1 I like this one - for example "please enter the word spelled by the third underlined red letter, fourth bold green letter, and fifth non-bold blue letter". – [Andrew](#) Sep 23, 2008 at 0:13

9 That example above
'ssdfatwerweajhcsadkoghvefdhrffghlfgdhowfgh' could be solved by a simple regex – [alex](#) Oct 29, 2008 at 1:48

This would not be good for users with acalcula. There are scientists with this affliction so it isn't unreasonable that there could be programmers with it. – [BobbyShaftoe](#) Dec 23, 2008 at 2:54

Excellent idea! Perhaps even by playing with changing foreground/background colors, you can get something that displays text easily visible to humans, but too random for bots? Of course this is harder on color-blind people :-(
– [Shalom Craimer](#) Jan 21, 2009 at 7:45

2 Using colour might complicate things, because you would need to support different forms of colour-blindness. Otherwise this sounds pretty good. – [KarstenF](#) Oct 25, 2010 at 11:28

12 Although this [similar discussion](#) was started:

votes



We are trying this solution on one of our frequently data mined applications:



[A Better CAPTCHA Control \(Look Ma - NO IMAGE!\)](#)

You can see it in action on our [Building Inspections Search](#).

You can view Source and see that the CAPTCHA is just HTML.

Share

edited May 23, 2017 at 12:03

community wiki

[2 revs](#)

[KP.](#)

That will work for NOW, but as soon as enough sites use an approach like that, spammers will render the html to an image and OCR the result. – [warp](#) Oct 25, 2008 at 8:36

12

votes

I know that no one will read this, but what about the **dog or cat** CAPTCHA?



You need to say which one is a cat or a dog, machines



can't do this.. <http://research.microsoft.com/asirra/>

Is a cool one..

Share

answered Nov 13, 2008 at 13:41

community wiki

[José Leal](#)

4 Computers can't do this. Nor can blind people. – [TRiG](#) Nov 9, 2010 at 14:54

1 Nor me, when looking at it on a high res monitor. Those images are TINY. Also one of them was of some unidentifiable creature behind a white fence. – [jsims281](#) Dec 14, 2010 at 17:36

10 I just use simple questions that anyone can answer:

votes



What color is the sky?



What color is an orange?

What color is grass?

It makes it so that someone has to custom program a bot to your site, which probably isn't worth the effort. If they do, you just change the questions.

Share

answered [Aug 22, 2008 at 16:06](#)

community wiki
[tghw](#)

Cyc can solve this trivially... and it's open source. Would require at most a couple hours of scripting to implement.
– [rmeador](#) Dec 2, 2008 at 19:04

this is used by ubuntu forum, also. i like it, and the implementations of checks like "2 + 2 = ?" or "what is the first letter of the alphabet" is very simple. – [pistacchio](#) Dec 29, 2008 at 0:26

8 The answers: 1) Right now, a light blue, later, red, then black with hints of orange near downtown. 2) orange, unless it's moldy, then it's green or black or white. 3) brown, in Southern California, unless you're in Beverly Hills, then it's green. – [mmr](#) Oct 21, 2009 at 22:35

52 @mmr See, that's actually a benefit of the system, it keeps the smartasses from posting comments... – [tghw](#) Oct 22, 2009 at 19:15

2 The second and third answers are biased towards people living in deserts or Baltimore. – [Mike Robinson](#) Jan 14, 2010 at 22:58

10
votes



I personally do not like CAPTCHA it harms usability and does not solve the security issue of making valid users invalid.

I prefer methods of bot detection that you can do server side. Since you have valid users (thanks to OpenID) you can block those who do not "behave", you just need to identify the patterns of a bot and match it to patterns of a typical user and calculate the difference.

Davies, N., Mehdi, Q., Gough, N. : Creating and Visualising an Intelligent NPC using Game Engines and AI Tools
<http://www.comp.glam.ac.uk/ASMTA2005/Proc/pdf/game-06.pdf>

Golle, P., Ducheneaut, N. : Preventing Bots from Playing Online Games <-- ACM Portal

Ducheneaut, N., Moore, R. : The Social Side of Gaming: A Study of Interaction Patterns in a Massively Multiplayer

Online Game

Sure most of these references point to video game bot detection, but that is because that was what the topic of our group's paper titled *Robot Wars: An In-Game Exploration of Robot Identification*. It was not published or anything, just something for a school project. I can email if you are interested. The fact is though that even if it is based on video game bot detection, you can generalize it to the web because there is a user attached to patterns of usage.

I do agree with MusiGenesis 's method of this approach because it is what I use on my website and it does work decently well. The invisible CAPTCHA process is a decent way of blocking most scripts, but that still does not prevent a script writer from reverse engineering your method and "faking" the values you are looking for in javascript.

I will say the best method is to 1) establish a user so that you can block when they are bad, 2) identify an algorithm that detects typical patterns vs. non-typical patterns of website usage and 3) block that user accordingly.

Share

answered [Feb 7, 2009 at 16:50](#)

community wiki
[jwendl](#)

Why can't a bot register OpenIDs? An attacker just needs to create their own OpenID publisher. – [rjmunro](#) Jul 24, 2009 at 17:49

Yes @rjmunro, and that is a good thing. The difficulty of the internet is identification of anonymous users. If a bot registers an OpenID and you identify that OpenID user as a bot then you can shut it down. It is no longer anonymous. That doesn't prevent multiple registrations by the same provider, but then you can shut that provider down for allowing bots. The goal is to remove the anonymity of the internet as best you can.
– [jwendl](#) Oct 17, 2009 at 16:28

10 I have some ideas about that I like to share with you...

votes



First Idea to avoid OCR



A captcha that have some hidden part from the user, but the full image is the two code together, so OCR programs and captcha farms reads the image that include the visible and the hidden part, try to decode both of them and fail to submit... - I have all ready fix that one and work online.

<http://www.planethost.gr/IdeaWithHiddenPart.gif>

Second Idea to make it more easy

A page with many words that the human must select the right one. I have also create this one, is simple. The words

are clickable images, and the user must click on the right one.

<http://www.planethost.gr/ManyWords.gif>

Third Idea with out images

The same as previous, but with divs and texts or small icons. User must click only on correct one div/letter/image, what ever.

<http://www.planethost.gr/ArrayFromDivs.gif>

Final Idea - I call it CicleCaptcha

And one more my **CicleCaptcha**, the user must locate a point on an image. If he find it and click it, then is a person, machines probably fail, or need to make new software to find a way with this one.

<http://www.planethost.gr/CicleCaptcha.gif>

Any critics are welcome.

Share

edited Jan 20, 2012 at 0:27

community wiki

2 revs

Aristos

Re-captcha has text to speech for its words. It could tell you where to click. – [Robert P](#) Dec 21, 2010 at 18:59

@RobertP Clicking when you're visually impaired is about like shooting when you're visually impaired... "Just aim for the head". – user166390 Jan 20, 2012 at 19:28

8 @pst "A little to the left. That's right, a little more. Good. Good. Now back. Up. No, up the other way. Yes, you've got it. Just a little more. Yes. Just a little more. Almost there. Almost there...almost...almost...CAPTCHA!!!" – [Robert P](#) Jan 21, 2012 at 6:22

9 [Best captcha ever!](#) Maybe you need something like this for sign-up to keep the riff-raff out.

votes



Share

answered [Aug 21, 2008 at 12:16](#)



community wiki
[Chris Bartow](#)

7 Recently, I started adding a tag with the name and id set to "message". I set it to hidden with CSS (display:none). Spam bots see it, fill it in and submit the form. Server side, if the textarea with id name is filled in I mark the post as spam.

votes



Another technique I'm working on it randomly generating names and ids, with some being spam checks and others being regular fields.

This works very well for me, and I've yet to receive any successful spam. However, I get far fewer visitors to my sites :)

Share

answered [Oct 28, 2010 at 1:21](#)

community wiki

[Brandon - Free Palestine](#)

Using css to hide the form field and asserting it is empty has worked for me as well. Not fool proof but is a good option.

– [Chris](#) Oct 28, 2010 at 13:03

Technique 1: [Honeypot](#) – [Kevin Ji](#) Aug 29, 2011 at 17:26 

6

votes



Very simple arithmetic is good. Blind people will be able to answer. (But as Jarod said, beware of operator precedence.) I gather someone could write a parser, but it makes the spamming more costly.

Sufficiently simple, and it will be not difficult to code around it. I see two threats here:

1. random spambots and the human spambots that might back them up; and
2. bots created to game Stack Overflow

With simple arithmetics, you might beat off threat #1, but not threat #2.


Share

edited Apr 19, 2010 at 16:18

community wiki
2 revs, 2 users 88%
Josh

A parser, I'd assume, is significantly easier than writing an image-captcha cracker. Remember, the easiest thing you offer to users is what a spambot will probably use. Sadly, the no-JS captcha needs to be harder. – [stalepretzel](#) Oct 1, 2008 at 11:58

5
votes



I've had amazingly good results with a simple "Leave this field blank:" field. Bots seem to fill in everything, particularly if you name the field something like "URL". Combined with strict referrer checking, I've not had a bot get past it yet.

Please don't forget about accessibility here. Captchas are notoriously unusable for many people using screen readers. Simple math problems, or very trivial trivia (I liked the "what color is the sky" question) are much more friendly to vision-impaired users.

Share

answered Sep 9, 2008 at 17:36

community wiki

5

votes



Simple text sounds great. **Bribe the community to do the work!** If you believe, as I do, that SO rep points measure a user's commitment to helping the site succeed, it is completely reasonable to offer reputation points to help protect the site from spammers.

Offer +10 reputation for each contribution of a simple question and a set of correct answers. The question should suitably far away (edit distance) from all existing questions, and the reputation (and the question) should gradually disappear if people can't answer it. Let's say if the failure rate on correct answers is more than 20%, then the submitter loses one reputation point per incorrect answer, up to a maximum of 15. So if you submit a bad question, you get +10 now but eventually you will net -5. Or maybe it makes sense to ask a sample of users to vote on whether the captcha question is a good one.

Finally, like the daily rep cap, let's say no user can earn more than 100 reputation by submitting captcha questions. This is a reasonable restriction on the weight given to such contributions, and it also may help prevent spammers from seeding questions into the system. For example, you could choose questions not with equal probability but with a probability proportional to the submitter's reputation. Jon Skeet, please don't submit any questions :-)

5

votes



Make an AJAX query for a cryptographic nonce to the server. The server sends back a JSON response containing the nonce, and also sets a cookie containing the nonce value. Calculate the SHA1 hash of the nonce in JavaScript, copy the value into a hidden field. When the user POSTs the form, they now send the cookie back with the nonce value. Calculate the SHA1 hash of the nonce from the cookie, compare to the value in the hidden field, and verify that you generated that nonce in the last 15 minutes (memcached is good for this). If all those checks pass, post the comment.

This technique requires that the spammer sits down and figures out what's going on, and once they do, they still have to fire off multiple requests and maintain cookie state to get a comment through. Plus they only ever see the `Set-Cookie` header if they parse and execute the JavaScript in the first place and make the AJAX request. This is far, far more work than most spammers are willing to go through, especially since the work only applies to a single site. The biggest downside is that anyone with JavaScript off or cookies disabled gets marked as potential spam. Which means that moderation queues are still a good idea.

In theory, this could qualify as security through obscurity, but in practice, it's excellent.

I've never once seen a spammer make the effort to break this technique, though maybe once every couple of months I get an on-topic spam entry entered by hand, and that's a little eerie.

Share

edited Oct 21, 2009 at 22:35

community wiki

2 revs

Bob Aman

5 1) Human solvers

votes



All mentioned here solutions are circumvented by human solvers approach. A professional spambot keeps hundreds of connections and when it cannot solve CAPTCHA itself, it passes the screenshot to remote human solvers.

I frequently read that human solvers of CAPTCHAs break the laws. Well, this is written by those who do not know how this (spamming) industry works.

Human solvers do not directly interact with sites which CAPTCHAs they solve. They even do not know from which sites CAPTCHAs were taken and sent them. I am aware about dozens (if not hundreds) companies or/and websites offering human solvers services but not a single one for direct interaction with boards being broken.

The latter do not infringe any law, so CAPTCHA solving is completely legal (and officially registered) business companies. They do not have criminal intentions and might, for example, have been used for remote testing, investigations, concept proofing, prototyping, etc.

2) Context-based Spam

AI (Artificial Intelligent) bots determine contexts and maintain context sensitive dialogues at different times from different IP addresses (of different countries). Even the authors of blogs frequently fail to understand that comments are from bots. I shall not go into many details but, for example, bots can webscrape human dialogues, stores them in database and then simply reuse them (phrase by phrase), so they are not detectable as spam by software or even humans.

[The most voted answer](#) telling:

- *The theory being that:
 - A spam bot will not support JavaScript and will submit what it sees
 - If the bot does support JavaScript it will submit the form instantly
 - The commenter has at least read some of the page before posting"

as well [honeypot answer](#) and most answers in this thread are just plain wrong.

I daresay they are **victim-doomed approaches**

Most spambots work through local and remote javascript-aware (patched and managed) browsers from different IPs (of different countries) and they are quite clever to circumvent honey traps and honey pots.

The different problem is that even blog owners cannot frequently detect that comments are from bot since they are really from human dialogs and comments harvested from other web boards (forums, blog comments, etc)

3) Conceptually New Approach

Sorry, I removed this part as precipitated one

Share

edited May 23, 2017 at 12:34

community wiki

11 revs

Gennady Vanin Геннадий
Ванин

-
- 1 You've outlined the problem with almost every answer on this thread. They all would be defeated quickly if they weren't exclusive to tiny website. If any of them were used on a large website or a number of small websites (say as a Wordpress plugin), they would be defeated in a day. They really aren't CAPTCHAS, but lucky cases of security through obscurity. You are also right on; modern spam is putting comments that even I cannot recognize as spam. I've given up on CAPTCHA and

instead use Mollom. Crowdsourcing spam is a better approach imho. – [Dan](#) Sep 12, 2011 at 22:44

5
votes



Actually it could be an idea to have a programming related captcha set. For example:

Language for Captcha Problem:

How many syntax errors are shown below:

```
$var == array(1 = 'one');;
```

Answer:


There is the possibility of someone building a syntax checker to bypass this but it's a lot more work to bypass a captcha. You get the idea of having a related captcha though.

Share

[edited Mar 29, 2012 at 18:05](#)

community wiki
[2 revs, 2 users 86%](#)
[Ross](#)

-
- 5 Except: Would you know that the answer to the example question (in case someone stumbles across this later and the image isn't there: "How many PHP syntax errors in `$var == array(1 = 'one');` is 1, not 3 ? (Seriously. Try `$var`

`== array(1 => 'one');;) ^_~` – [pinkgothic](#) Nov 8, 2010 at 10:09 

Heh, it's been long enough that I don't remember making that image. You're right of course, and I don't think this would be very good as a captcha - but maybe as an entry requirement for a programming forum. – [Ross](#) Nov 8, 2010 at 23:17

4
votes

What if you used a combination of the captcha ideas you had (choose any of them - or select one of them randomly):



- ASCII text captcha: `//(_)//`
- math puzzles: what is 7 minus 3 times 2?
- trivia questions: what tastes better, a toad or a popsicle?

with the addition of placing the exact same captcha in a css hidden section of the page - the honeypot idea. That way, you'd have one place where you'd expect the correct answer and another where the answer should be unchanged.


Share

answered [Aug 25, 2008 at 19:42](#)

community wiki

[TheEmirOfGroofunkistan](#)

- 1 "What tastes better" is quite subjective, in my opinion. People who go out on a limb concerning taste will be interpreted as a

bot. Moreover, with only two answers the chance of passing as a bot is 50%. – [pimvdb](#) Jun 18, 2011 at 10:43 

True on the subjective part, but the particular was from the original question. My point was to use all three (or multiple) types randomly. Also, the chance of passing is only 50% if the answer is multiple choice. If you have the user type the word without indicating the possible answers in a list, it becomes much more difficult for a bot to pick out the correct words in the question/answer. – [TheEmirOfGroofunkistan](#) Jul 12, 2011 at 14:33

4

votes



I have to admit that I have no experience fighting spambots and don't really know how sophisticated they are. That said, I don't see anything in the jQuery article that couldn't be accomplished purely on the server.

To rephrase the summary from the jQuery article:

1. When generating the contact form on the server ...
2. Grab the current time.
3. Combine that timestamp, plus a secret word, and generate a 32 character 'hash' and store it as a cookie on the visitor's browser.
4. Store the hash or 'token' timestamp in a hidden form tag.
5. When the form is posted back, the value of the timestamp will be compared to the 32 character 'token' stored in the cookie.

6. If the information doesn't match, or is missing, or if the timestamp is too old, stop execution of the request ...

Another option, if you want to use the traditional image CAPTCHA without the overhead of generating them on every request is to pre-generate them offline. Then you just need to randomly choose one to display with each form.

Share

answered [Sep 16, 2008 at 16:41](#)

community wiki
[Peter](#)

1

2

3

4

Next