When the bots attack! [closed]

Asked 16 years, 3 months ago Modified 4 years, 1 month ago

Viewed 4k times

26

As it currently stands, this question is not a good fit for our Q&A format. We expect answers to be supported by facts, references, or expertise, but this question will likely solicit debate, arguments, polling, or extended discussion. If you feel that this question can be improved and possibly reopened, visit the help center for guidance.

Closed 12 years ago.

What are some popular spam prevention methods besides CAPTCHA?

security captcha spam-prevention bots

Share

Improve this question

Follow

edited Feb 23, 2010 at 11:47

AviD

13.1k ● 7 ● 64 ● 93

asked Sep 21, 2008 at 18:13

Jose Vega

10.2k ● 7 ● 42 ● 58

1 Looks like almost all the responders missed that you said BESIDES captcha... Which is a great angle, since CAPTCHA is broken and does not work. – AviD Sep 21, 2008 at 20:31

27 Answers

Sorted by: Highest score (default)

23

I have tried doing 'honeypots' where you put a field and then hide it with CSS (marking it as 'leave blank' for anyone with stylesheets disabled) but I have found that a lot of bots are able to get past it very quickly. There are also techniques like setting fields to a certain value and changing them with JS, calculating times between load time and submit time, checking the referer URL, and a million other things. They all have their pitfalls and pretty much all you can hope for is to filter as much as you can with them while not alienating who you're here for: the users.

At the end of the day, though, if you really, really, don't want bots to be sending things through your form you're going to want to put a CAPTCHA on it - best one I've seen that takes care of mostly everything is reCAPTCHA - but thanks to India's CAPTCHA solving market and the ingenuity of spammers everywhere that's not even successful all of the time. I would beware using something that is 'ingenious' but kind of 'out there' as it would be more of a 'wtf' for users that are at least somewhat used to your usual CAPTCHAs.

Share Improve this answer Follow

edited Sep 21, 2008 at 18:35

answered Sep 21, 2008 at 18:18

Paolo Bergantino
488k ● 82 ● 521 ● 437

I like the CSS technique. It works very well across the board. I'd also vote for this answer, but I have no votes left! :D – Till Sep 21, 2008 at 22:17

1 Actually, this is trivially bypassable by the simplest of techniques; if a spammer wants to misuse your site, he's gonna. Only thing protecting you is if you're not big enouhg to bother with looking at your code. – AviD Sep 23, 2008 at 4:45

1 -1 This is funny that solutions routinely circumvented automatically by spam bots without even the need to engage human solvers shops APIs integrated into professional bots are marked as answer. Is it by ignorance or intentionally? – Gennady Vanin Геннадий Ванин Dec 21, 2010 at 17:34

18

Shocking, but almost every response here included some form of CAPTCHA. The OP wanted something different, I guess maybe he wanted something that actually works, and maybe even solves the real problem.

CAPTCHA **doesn't** work, and even if it did - its the wrong problem - humans can still flood your system, and by definition CAPTCHA wont stop that (cuz its designed only to tell if you're a human or not - not that it does that well...)

So, what other solutions *are* there? Well, it depends... on your system and your needs. For instance, if all you're trying to do is limit how many times a user can fill out a "Contact Me" form, you can simply throttle how many requests each user can submit per hour/day/whatever. If your users are anonymous, maybe you need to throttle according to IP addresses, and occasionally blacklist an IP (though this too can be circumvented, and causes other problems).
If you're referring to a forum or blog comments (such as this one), well the more I use it the more I like the solution. A mix between authenticated users, authorization (based on reputation, not likely to be accumulated through flooding), throttling (how many you can do a day), the occasional CAPTCHA, and finally community moderation to cleanup the few that get through - all combine to provide a decent solution. (I wonder if Jeff can provide some info on how much spam and other malposts actually get through...?)

Another control to consider (dont know if they have it here), is some form of IDS/IPS - if you can detect and recognize spam, you can block THAT pattern. Moderation fills that need manually, here...

Note that any one of these does not *prevent* the spam, but incrementally *lowers the probability*, and thus the profitability. This changes the economic equation, and leaves CAPTCHA to actually provide enough value to be worth it - since its no longer worth it for the spammers to

bother breaking it or going around it (thanks to the other controls).

Share Improve this answer

Follow

answered Sep 21, 2008 at 20:42

AviD

13.1k ● 7 ● 64 ● 93

12

Give the user the possibility to calculate:

What is the sum of 3 and 8?

By the way: Just surfed by an interesting approach of Microsoft Research: Asirra.

http://research.microsoft.com/asirra/

It shows you several pictures and you have to identify the pictures with a given motif.

Share Improve this answer

Follow

answered Sep 21, 2008 at 18:15

Johannes Hädrich

1,223 ● 1 ● 12 ● 20

1 I have used this one to great effect in the past as well. What is 2+2? Anything that is custom, easy for users, and easy for you is the best solution. – Jason Short Sep 23, 2008 at 2:58

CAPTCHA means automated turing test, and asking humans quesions falls under that definition. – Kornel Oct 19, 2008 at 11:50

I checked out Aksimet, found it to be very interesting, and a great idea since animal recognition is something bots are

horrible at. I'd upvote 5 times if I could based on this alone :D – David Frenkel Nov 13, 2008 at 20:13

1 if onley i'd bought $20 worth of MS in 95 :/ – divinci Jun 7, 2009 at 0:55

I've used Asirra on my site, and it has worked fantastically. I recommend this to anybody. Also, it's not frustrating like CAPTCHA is. It's fun for the user and helps pets get adopted. Win for me, win for my users, win for abandoned pets. – Eric Jun 18, 2009 at 16:17

10

Try Akismet

Captchas or any form of human-only questions are horrible from a usability perspective. Sometimes they're necessary, but I prefer to kill spam using filters like Akismet.

Akismet was originally built to thwart spam comments on WordPress blogs, but the API is capabable of being adapted for other uses.

Update: We've started using the ruby library Rakismet on our Rails app, Yarp.com. So far, it's been working great to thwart the spam bots.

Share Improve this answer

Follow

edited Nov 17, 2020 at 23:00

Hirurg103

4,953 ● 2 ● 39 ● 53

answered Sep 21, 2008 at 21:33

Ryan McGeary

1 I'd be interested to know just what they base it on, but I suppose it's mostly common patterns liek advertising combined with URL link – David Frenkel Nov 13, 2008 at 20:15

I think they use a number of inputs, including content against a bayesian-like filter, URLs, source IP addresses, etc. – Ryan McGeary Nov 14, 2008 at 18:54

7

A very simple method which puts no load on the user is just to disable the submit button for a second after the page has been loaded. I used it on a public forum which had continuous spam posts, and it stopped them since.

Share Improve this answer

Follow

answered Sep 21, 2008 at 19:59

Pietro Polsinelli

Pretty interesting suggestion, i might try that out. Thanks for your answer. – Jose Vega Sep 21, 2008 at 22:34

2 I don't get it. Do the bots care about whether the submit button is enabled? – Seun Osewa Feb 19, 2009 at 0:31

I guess they click it while it's disabled, revealing the fact that they don't respect the timer. – Paweł Polewicz Dec 21, 2009 at 16:25

-1 - bots almost always ignore javascript. – Lotus Notes May 21, 2010 at 18:46

7

Ned Batchelder wrote up a technique that combines hashes with honeypots for some wickedly effective bot-prevention. No captchas, just code.

It's up at [Stopping spambots with hashes and honeypots]:

> Rather than stopping bots by having people identify themselves, we can stop the bots by making it difficult for them to make a successful post, or by having them inadvertently identify themselves as bots. This removes the burden from people, and leaves the comment form free of visible anti-spam measures.
>
> This technique is how I prevent spambots on this site. It works. The method described here doesn't look at the content at all. It can be augmented with content-based prevention such as Akismet, but I find it works very well all by itself.

Share Improve this answer

Follow

edited Jun 20, 2020 at 9:12

Community Bot

1 ● 1

answered Oct 1, 2008 at 2:58

joemurphy

608 ● 4 ● 10

[http://chongqed.org/] maintains blacklists of active spam sources and the URLs being advertised in the spams. I

5

have found filtering posts for the latter to be very effective in forums.

Share Improve this answer

Follow

answered Sep 21, 2008 at 18:55

moonshadow

88.9k ● 7 ● 86 ● 121

4

The most common ones I've observed orient around user input to solve simple puzzles e.g. of the following is a picture of a cat. (displaying pictures of thumbnails of dogs surrounding a cat). Or simple math problems.

While interesting I'm sure the arms race will also overwhelm those systems too.

Share Improve this answer

Follow

answered Sep 21, 2008 at 18:16

stephbu

5,082 ● 28 ● 42

4

You can use Recaptcha to at least make a captcha useful. Then you can make questions with simple verbal math problems or similar. Microsoft's Asirra makes you find pics of cats and dogs. Requiring a valid email address to activate an account stops spammers when they wouldn't get enough benefit from the service, but might deter normal users as well.

Share Improve this answer

Follow

answered Sep 21, 2008 at 18:19

jjrv

4,335 ● 3 ● 43 ● 55

3

The following is unfeasible with today's technology, but I don't think it's too far off. It's also probably overkill for dealing with forum spam, but could be useful for account sign-ups, or any situation where you wanted to be really sure you were dealing with humans and they would be prepared for it to take a few minutes to complete the process.

Have 2 users who are trying to prove themselves human connect to each other via their webcams and ask them if the person they are seeing is human and live (i.e. not a recording), by getting them to, for example, mirror each other's movements, or write something on a piece of paper. Get everyone to do this a few times with different users, and throw a few recordings into the mix which they also have to identify correctly as such.

Share Improve this answer

Follow

answered Sep 24, 2008 at 1:57

Sam Hasler

12.6k ● 10 ● 73 ● 106

+1 for creativity and thinking outside the box =) ...beware of the 'think of the children' crowd though, if you implement it... – David Thomas Jun 18, 2009 at 17:03

3

A popular method on forums is to simply queue the threads of members with less than 10 posts in a moderation queue. Of course, this doesn't help if you don't have moderators, or it's not a forum. A more general method is the calculation of hyperlink to text ratios. Often, spam posts contain a ton of hyperlinks, and you can catch a lot this way. In the same vein is comparing the content of consecutive posts. Simply do not allow consecutive posts that are extremely similar.

Of course, anyone with knowledge of the measures you take is going to be able to get around them. To be honest, there is little you can do if you are the target of a specific attack. Rather, you should focus on preventing more general, unskilled attacks.

Share Improve this answer

Follow

answered Sep 24, 2008 at 2:04

user7545

3,060 ● 4 ● 23 ● 22

3

For human moderators it surely helps to be able to easily find and delete all posts from some IP, or all posts from some user if the bot is smart enough to use a registered account. Likewise the option to easily block IP addresses or accounts for some time, without further administration, will lessen the administrative burden for human moderators.

Using cookies to make bots and human spammers believe that their post is actually visible (while only they

themselves see it) prevents them (or trolls) from changing techniques. Let the spammers and trolls see the other spam and troll messages.

Share Improve this answer

Follow

answered Jun 18, 2009 at 16:12

Arjan

23.5k ● 12 ● 62 ● 71

2

Javascript evaluation techniques like this Invisible Captcha system require the browser to evaluate Javascript before the page submission will be accepted. It falls back nicely when the user doesn't have Javascript enabled by just displaying a conventional CAPTCHA test.

Share Improve this answer

Follow

answered Sep 21, 2008 at 18:15

Jon Galloway

53.1k ● 25 ● 127 ● 194

There are a bunch of bots out there now with script execution capabilities. I think the days of using script as a gatekeeper are numbered. – stephbu Sep 21, 2008 at 18:18

Yeah, they soon will be executing JS faster than most browsers, thanks Google :) – Ilya Ryzhenkov Sep 21, 2008 at 19:38

Indeed JavaScript is a sort of poor gatekeeper. – Till Sep 21, 2008 at 22:17

Animated captchas' - scrolling text - still easy to recognize by humans but if you make sure that none of the frames

2

offer something complete to recognize.

multiple choice question - All it takes is a _____ and a smile. idea here is that the user will have to choose/understand.

session variable - checking that a variable you put into a session is part of the request. will foil the dumb bots that simply generate requests but probably not the bots that are modeled like a browser.

math question - 2 + 5 = - this again is to ask a question that is easy to solve but prevents the bots ability to generate a response.

image grid - you create grid of images - select 1 or 2 of a particular type such as 3x3 grid picture of animals and you have to pick out all the birds on the grid.

Hope this gives you some ideas for your new solution.

Share Improve this answer

Follow

answered Sep 21, 2008 at 18:41

MikeJ

14.6k ● 22 ● 74 ● 89

I see math questions getting more and more common, but they don't really seem like the best choice. When they become prevalent enough to annoy botters, it will probably be very simple to add a simple function to parse the trivial arithmetic that most of these questions are. – Jeremy Banks Sep 21, 2008 at 22:18

A single multiple choice question with N options isn't very effective since it will still allow a random guess to succeed 1 /

N of the time. The beauty of "pick out all the birds in the 3x3 grid" is that 1 / 2^9 << 1 / 4 (wlog taking 4 as the usual number of answers to a multiple choice question).
– Doug McClean Jun 18, 2009 at 16:08

2

A friend has the simplest anti-spam method, and it works.

He has a custom text box which says "please type in the number 4".

His blog is rather popular, but still not popular enough for bots to figure it out (yet).

Share Improve this answer

Follow

answered Sep 21, 2008 at 19:06

ripper234

230k ● 280 ● 642 ● 913

3 I hope that number was randomly chosen! xkcd.com/221
– Loofer Sep 21, 2008 at 21:11

Aidan, I had the same thought! – epochwolf Oct 1, 2008 at 3:10

2

Please remember to make your solution accessible to those not using conventional browsers. The iPhone crowd are not to be ignored, and those with vision and cognitive problems should not be excluded either.

Share Improve this answer

Follow

answered Sep 21, 2008 at 21:14

Loofer

6,969 ● 9 ● 65 ● 106

2

Honeypots are one effective method. Phil Haack gives one good honeypot method, that could be used in principle for any forum/blog/etc.

You could also write a crawler that follows spam links and analyzes their page to see if it's a genuine link or not. The most obvious would be pages with an exact copy of your content, but you could pick out other indicators.

Moderation and blacklisting, especially with plugins like these ones for WordPress (or whatever you're using, similar software is available for most platforms), will work in a low-volume environment. If your environment is a low volume one, don't underestimate the advantage this gives you. Personally deciding what is reasonable content and what isn't gives you ultimate flexibility in spam control, if you have the time.

Don't forget, as others have pointed out, that CAPTCHAs are not limited to text recognition from an image. Visual association, math problems, and other non-subjective questions relayed through an image also qualify.

Share Improve this answer

Follow

edited Sep 23, 2008 at 2:12

answered Sep 21, 2008 at 18:22

Dustman

5,263 ● 10 ● 34 ● 40

2

[Sblam](#) is an interesting project.

Share Improve this answer

Follow

answered Jun 18, 2009 at 16:05

Michal M

9,470 ● 8 ● 50 ● 64

1

Invisble form fields. Make a form field that doesn't appear on the screen to the user. using display: none as a css style so that it doesn't show up. For accessibility's sake, you could even put hidden text so that people using screen readers would know not to fill it in. Bots almost always fill in all fields, so you could block any post that filled in the invisible field.

Share Improve this answer

Follow

answered Sep 21, 2008 at 18:18

Kibbee

66.1k ● 28 ● 144 ● 184

interesting! do you have experience with this method in production use? – Johannes Hädrich Sep 21, 2008 at 18:20

It wasn't very successful the one time I tried it. The website is still live employing that technique and no matter what variation of it I do the bots get through. Maybe I was just

targeted by the smarter ones =/ – Paolo Bergantino Sep 21, 2008 at 18:44

1

Block access based on a blacklist of spammers IP addresses.

Share Improve this answer

Follow

answered Sep 21, 2008 at 18:19

Chris

577 ● 3 ● 14

This technique does not work at all and blacklists will usually contain mostly IPs of legitimate users. Almost anyone you succeed in blocking this way will be a legitimate user. – MarkR Sep 21, 2008 at 18:21

"This technique does not work at all" seems a bit of an overzealous claim to me. Blocking by IP address can root out some of the most serious offenders. I was not suggesting deny access to large blocks of addresses. – Chris Sep 21, 2008 at 18:31

IP blacklists need to be incrementally temporary, meaning it starts off temporary (to prevent blocking legitimate users), and each repetition it should stay on the blacklist longer. – AviD Sep 21, 2008 at 20:46

1

Honeypot techniques put an invisible decoy form at the top of the page. Users don't see it and submit the correct form, bots submit the wrong form which does nothing or bans their IP.

Share Improve this answer

Follow

answered Sep 21, 2008 at 18:20

Jon Galloway

53.1k • 25 • 127 • 194

1

I've seen a few neat ideas along the lines of Asira which ask you to identify which pictures are cats. I believe the idea originated from KittenAuth a while ago..

Share Improve this answer

Follow

answered Sep 23, 2008 at 2:35

Jon Cage

37.4k • 41 • 145 • 222

1

Use something like the google image labeler with appropriately chosen images such that a computer wouldn't be able to recognise the dominant features of it that a human could.

The user would be shown an image and would have to type words associated with it. They would keep being shown images until they have typed enough words that agreed with what previous users had typed for the same image. Some images would be new ones that they weren't being tested against, but were included to record what words are associated with them. Depending on your audience you could also possibly choose images that only they would recognise.

Share Improve this answer

Follow

answered Sep 24, 2008 at 1:57

Sam Hasler

12.6k ● 10 ● 73 ● 106

1

[Mollom](#) is supposedly good at stopping spam. Both personal (free) and professional versions are available.

Share Improve this answer

Follow

answered Nov 7, 2008 at 22:56

Pieter

17.7k ● 8 ● 53 ● 90

1

I know some people mentioned ASIRRA, but if you go to all the adopt me links for the images, it will say on that linked page if its a cat or dog. So it should be relatively easy for a bot to just go to all the adoptme links. So its just a matter of time for that project.

Share Improve this answer

Follow

answered Jan 16, 2009 at 16:17

tooleb

612 ● 3 ● 6 ● 14

1

just verify the email address and let google/yahoo etc worry about it

Share Improve this answer

Follow

answered Jun 18, 2009 at 15:55

MatthewFord

2,926 ● 2 ● 21 ● 32

-2

You could get some device ID software the41 has some fraud prevention software that can detect the hardware being used to access your site. I belive they use it to catch fraudsters but could be used to stop bots. Once you have identified an device being used by a bot you can just block that device. Last time a checked it can even trace your route throught he phone network (Not your Geo-IP !!) so can even block a post code if you want.

Its expensive through so prop. a better cheaper solution that is a little less big brother.

Share Improve this answer

Follow

answered Sep 23, 2008 at 13:36

justin Case