# Captcha

Согласно своему определению, captcha — это автоматизированный публичный тест Тьюринга (тест который может пройти человек, но не компьютер). В статье я буду рассматривать свойтсва капчи на примере самого распространненого её вида — текста на картинке, хотя почти все написанное одинаково применимо к любому виду капчи.

## Два главных свойства капчи

Любая капча должна обладать двумя свойствами, без которых она не будет работать:

Устойчивость к распознаванию — свойство, защищающее капчу от распознавания алгоритмом — например системой распознавания текста. Гарантирует то, что человек сможет прочитать текст на картинке, а компьютер нет.

Антипример: стандартная капча форумов phpBB 2.x таким свойством не обладала — из-за относительной простоты распознавания появились скрипты, которые спамили все подряд форумы вынуждая веб-мастеров менять капчу на более стойкую.

Устойчивость к угадыванию — свойство капчи, не позволяющее угадать её значение за небольшое число попыток (менее 1000). Если набор возможных значений капчи невелик, программе не составит труда угадать её подбором вместо распознавания.

Антипример: арифметическая капча вроде «1+2» (перебор чисел от 1 до 20 в скором времени даст результат).

Антипример: выбрать из нескольких картинок ту, на которой изображён котик.

## Проверка капчи

Значение для проверки должно храниться на сервере, а не передаваться вместе с картинкой в браузер. Для сопоставления посетителя и правильного значения капчи необходимо использовать некий ключ, который передаётся вместе с капчей (идентификатор сессии, номер капчи и т.п.)

Антипример: если передавать саму капчу и значение для ее проверки (в том числе зашифрованное), то человеку достаточно один раз распознать такую капчу и далее использовать комбинацию «ответ»-«значение для проверки» в своём скрипте (по ссылке в начале поста как раз такой случай)

Перед проверкой ответа — надо убедиться, что он не пустой. В противном случае, злоумышленник может, не загружая картинку или удалив идентификатор текущей сессии, передать пустое значение и пройти капчу, т.к. произойдёт сравнение двух пустых строк (в PHP несуществующее значение равно пустой строке).

Антипример: уже упомянутый мной код if($\_POST['captcha'] == $\_SESSION['captcha']) return true;

Причем этот код был написан опытным программистом.

После проверки, сохраненное значение капчи необходимо удалить. Если не сделать этого, злоумышленник сможет использовать данное значение снова неограниченное число раз. Да, при обновлении страницы с формой обновляется и капча (либо при генерации формы, либо при генерации картинки), вот только скрипт может не загружать форму снова (надо упомянуть, что это не актуально если на сайте используются одноразовые csrf-токены для форм).

Антипример: гипотетическая форма логина, в которой достаточно один раз ввести капчу правильно, и далее подбирать пароль скриптом, избегая перегенерации капчи на сервере.

## Пуленепробиваемая капча

Защита от перебора. Если ваша капча устойчива к распознаванию, но не очень устойчива к перебору (например на ней надо прочитать всего 3-4 цифры), желательно ограничить число неправильных ответов «с одного ip» / «для одного логина» / etc. Такие ограничения необходимо проверять ДО проверки самой капчи (то есть даже в случае правильно введенной капчи, при наличии ограничения она не должна считаться пройденной) иначе оно не будет препятствовать перебору.

Защита от DoS. При генерации капчи на своем сервере, надо понимать, что это удобный вектор проведения DoS атак (которую, в отличие от DDoS, может устроить любой школьник). Для защиты можно ограничить число генерации капчи для одного ip, кэшированием капч и т.д.

Защита от распознавания. Если вы выбираете капчу, или вдруг собираетесь написать её сами, желательно понимать какая капча более защищена от распознавания. Существуют готовые универсальные скрипты распознавания капчи, работающие по принципу OCR, а в случае, если ваш сайт заинтересует спамеров есть риск, что будут использовать / писать скрипт конкретно под вашу капчу. Последнее правда относится больше к сайтам уровня Яндекс или vk, а вот вариант с защитой от банальных OCR желательно предусмотреть.

Защита от антигейтов. Если говорить формально, то капча как тест Тьюринга не обязана защищать вас от антигейтов, так как в этом случае её будет распознавать человек. С практической же точки зрения, этот вопрос весьма актуален и защищаться как-то надо.

Тут нет и не может быть «золотого стандарта» (ибо в таком случае антигейты внедрят его поддержку), поэтому вы вольны дополнять капчу любыми ухищрениями, чтобы сделать её распознавание через антигейт невозможным. Например:

— нестандартная капча (сбор паззла, поворот изображения, клик по области на фото и т.п.);

— кириллическая капча — самое простое решение, но имеет ряд минусов: подходит только для проектов с русскоязычной аудиторией, есть антигейты с поддержкой кириллицы;

— использование виртуальной клавиатуры рядом с капчей для ввода нестандартных символов или фигур (может быть неудобно пользователям мобильных);

## Юзабилити

Не просите ввести капчу, если вы уже убедились, что перед вами человек. Тут, однако, надо быть осторожным, чтобы форму нельзя было использовать скриптом неограниченное число раз после однократного ввода капчи человеком.

Пример: форма регистрации. Если я где-то регистрируюсь, и забыл ввести поле «почтовый индекс», но правильно ввёл капчу — не надо показывать мне новую. Потратьте 10 минут на то, чтобы сохранить где-то у себя, что вот эту конкретную форму сейчас пытается заполнить живой человек.

Для облегчения распознавания человеком: не используйте в капче одновременно буквы и цифры, не используйте одновременно прописные и строчные буквы, исключите похожие символы.

## Отказ от использования капчи

Лучшая капча — отсутствие капчи. Там, где можно отказаться от её использования — это надо сделать. Возможно, для этого придется реализовать дополнительные лимиты и проверки, но пользователи скажут вам спасибо.

Но тут надо быть очень осторожным. Например: форма регистрации без капчи, с полем email на который приходит письмо с активацией. Без дополнительных средств защиты такую форму могут завалить «левыми» адресами, и ваш сайт включат в черные списки почтовые службы. В таком случае можно обходиться без капчи, но только если у вас есть другой рубеж защиты, вроде лимита по ip.

## Источники

* <https://ru.wikipedia.org/wiki/Капча>
* <https://habr.com/ru/post/175461/>
* <https://myblaze.ru/vvod-kapchi/>
* <https://habr.com/ru/post/121010/>