

CAPTCHA ANALYSIS AND ITS PROBLEMS

M. Iavich, G. Iashvili

Scientific Cyber Security Association (SCSA)

ABSTRACT

CAPTCHA is a tool for recognizing a robot when trying to register, or filling out any other form on the site. CAPTCHA stands for - Completely Automated Public Turing test to recognize computer among the users. In this article we analyze Captcha and NoCaptcha. It is shown that for an ordinary user of the site NoCAPTCHA is more convenient, and intuitive than CAPTCHA. Considering the fact that the main problem for any CAPTCHA is its implementation, we show that the mechanism of protection from bots in any CAPTCHA is based on the actions of an ordinary user, and in this situation the user acts as a filter, what is not entirely fair, moreover, the latter conducts for the system work, for which he certainly is not paid. We consider that developers definitely need to consider the needs, and what is more important, the capabilities of average users. Based on this, to construct first of all convenient, and effective defense mechanism. Convenience of consumption is a priority for any project, and the loss of potential users due to obstacles in the form of a strong complication of the registration process, which is currently quite relevant - is a losing option for any project.

KEYWORDS: analysis, CAPTCHA, NoCAPTCHA

АННОТАЦИЯ

САРТЧА является инструментом для распознавания робота при попытке регистрации, либо заполнения любой другой форму на сайте. Расшифровывается САРТЧА как- Completely Automated Public Turing test to tell Computers and Humans Apart - полностью автоматизированный публичный тест Тьюринга для выявления компьютера среди посетителей сайта. В данной статье мы проводим анализ Captcha и NoCaptcha. Показано, что, что No САРТЧА более удобна, и интуитивно понятна для рядового пользователя сайта, чем САРТЧА. Исходя из того, что главная проблема для любой САРТЧА на сегодняшний день - это ее исполнение показано, что механизм защиты от ботов в любой САРТЧА завязан на действиях рядового пользователя, и в данной ситуации пользователь выступает в роли фильтра, что само по себе не совсем справедливо, к тому же, последний проводит для системы работу, за которую ему конечно же никто не платит.

Мы считаем, что разработчикам безусловно надо учитывать потребности, и что более важно, возможности среднестатистических пользователей. И основываясь на этом конструировать в первую очередь удобный, и эффективный механизм защиты. Удобство потребления является приоритетом для любого проекта, и потеря потенциальных пользователей из-за преград в виде сильного усложнения процесса регистрации, что на сегодняшний день достаточно актуально - является проигрышным вариантом для любого проекта.

CAPTCHA является инструментом для распознавания робота при попытке регистрации, либо заполнения любой другой формы на сайте. Расшифровывается CAPTCHA как - Completely Automated Public Turing test to tell Computers and Humans Apart - полностью автоматизированный публичный тест Тьюринга для выявления компьютера среди посетителей сайта.

В интернете существует большое количество способов заработка, и далеко не все из них законные. В данном случае мы имеем дело с крайне актуальной проблемой во всемирной паутине - спам рассылкой. Злоумышленники посредством различных интернет сайтов занимаются рассылкой спама, или другими словами “мусора”, который может содержать как рекламу, так и вредоносное программное обеспечение, либо схемы обмана получателей рассылки. Для этих целей создаются специальные программы - боты, которые вместо людей регистрируются на сайтах через соответствующие формы. Проходить регистрацию на сотнях сайтов человеку на порядок сложнее и займет данный процесс гораздо больше времени, боту же в свою очередь не нужен отдых, пища, либо сон, и регистрация даже на нескольких сотнях сайтов в день для него не составит особого труда.

Рано, или поздно практически любой пользователь глобальной сети интернет сталкивается с различными формами, будь то регистрация, или любая другая форма для заполнения. Помимо стандартных полей для заполнения, вроде электронной почты и имени пользователя, зачастую можно встретить дополнительную графу. Для завершения заполнения формы, пользователю нужно вписать сгенерированные системой символы, а иногда и слова. Символы в основном появляются перед пользователем в размытом посредством эффектов виде, это и есть CAPTCHA.

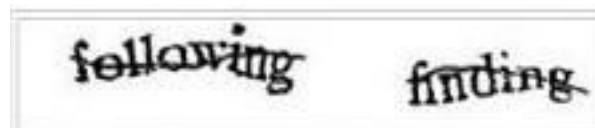


Рис. 1

Основываясь на условиях безопасности, и учитывая потребности и навыки обычных пользователей можно сформулировать несколько основных принципов для создания “идеальной” CAPTCHA. [1]

1. Прохождения CAPTCHA для человека не должно составлять труда, в то же время для робота не должно существовать способа обхода данного механизма.

К сожалению такой баланс в использовании данного защитного механизма достигается с трудом. Причиной тому постоянная работа по нахождению новых путей обхода CAPTCHA ботами. И это логично, ведь спамеры ни в коем случае не хотят отказываться от своей золотой жилы. В противовес данному процессу, ведется работа и по усложнению самой CAPTCHA. Удобство механизма защиты напрямую зависит от квалификации его разработчиков, и понимания ими психологии рядового пользователя. У последних часто возникают проблемы с распознаением символов в CAPTCHA.



Рис. 2

Большинство пользователей при виде данного изображения, передумают зарегистрироваться на вашем сайте, а то и хуже, в спешке его покинут.

Технология No CAPTCHA от корпорации Google работает по отличному от классической CAPTCHA механизму. Данная система оценивает поведение пользователя, а не его способность разгадывать отобразившиеся символы. Механизм No CAPTCHA работает следующим образом - посетитель сайта ставит галочку в графе "я не робот", после чего программа оценивает его поведение на признаки бота, к примеру время, проведенное на странице, IP адрес и прочие параметры. Если система заподозрит в пользователе робота, то к примеру, ему будет предложено отметить определенные объекты из появившихся на экране изображений.



Рис. 3

О самом механизме No CAPTCHA известно не много, разработчики не раскрывают деталей его работы, вскользь упоминая лишь элементы машинного обучения и глубокого анализа данных и рисков.

2. Удобность использования / понятный дизайн - после прохождения пользователем CAPTCHA хотя бы один раз, у него не должно возникнуть проблем с ее вводом впредь

Неопытного пользователя, CAPTCHA в классическом своем проявлении может отпугнуть непонятными, размытыми символами. Не всем сразу же будет понятно, что надо делать для завершения регистрации. На помощь пользователям в таком случае приходят подсказки в виде текста, он может встречаться в виде надписи перед, либо после поля, оповещающей пользователя о том, что он должен ввести соответствующие символы в поле. После прохождения данной процедуры, встретив идентичный тип защиты, пользователь легко в нем сориентируется.

В случае No CAPTCHA, пользователю надо лишь поставить галочку в поле "Я не робот", при анализе поведения пользователя, система может либо идентифицировать пользователя автоматически, либо предложит пользователю ряд изображений, из которых надо будет выбрать определенные объекты. Один клик для любого пользователя будет

легче, чем разбор символов и далее их ввод в соответствующее поле. Даже в случае отображения картинок, человеком они воспринимаются лучше, и процесс выделения нужных элементов занимает мало времени. [2]

3. Минимизировать ввод данных, для упрощения процесса регистрации

Если по данному параметру рассматривать классическую CAPTCHA, то в ней пользователю надо вводить данные, которые отобразились рядом с полем для их ввода. Символы как мы уже увидели в примере выше не всегда легко различить, и порой данный процесс связан с потерей дополнительного времени. Следовательно, процесс регистрации в целом займет больше времени.

В случае No CAPTCHA пользователь сначала отмечает поле CAPTCHA галочкой, затем система либо пропускает его, либо предлагает подтвердить его действия посредством выбора из ряда изображений. Пользователь при выборе неверных изображений получает новую порцию уже из новых картинок. Но, по нашему мнению, визуально объект на рисунке определить в разы проще нежели искаженные символы

4. Адаптированные типы CAPTCHA, как например аудио дорожка для людей со слабым зрением

В данном параметре оба вида CAPTCHA обладают дополнительной функцией проигрывания. И ей можно воспользоваться в случае, если пользователь не может разобрать изображения либо символы

Рассматривая систему No CAPTCHA вместе с классической CAPTCHA со стороны пользователя, можно сказать, что No CAPTCHA более удобна, и интуитивно понятна для рядового пользователя сайта. А в определенных случаях, после идентификации пользователя система автоматически ставит галочку в графе "я не робот". В свою очередь, обработка и распознавание данных с изображений в данной системе для бота гораздо труднее, нежели идентификация символов в классическом варианте CAPTCHA. Что теоретически делает данную систему, в ее нынешнем исполнении более эффективной против спам ботом по сравнению с классической версией.

Мы считаем, что самая главная проблема для любой CAPTCHA на сегодняшний день - это ее исполнение. По факту, спам рассылка, и боты являются проблемой для владельца сайта, и сервера, но никак не является проблемой для пользователей, которые на этом сайте регистрируются. А механизм защиты от ботов в любой CAPTCHA завязан на действиях рядового пользователя, и в данной ситуации пользователь выступает в роли фильтра, что

само по себе не совсем справедливо, к тому же, последний проводит для системы работу, за которую ему конечно же никто не платит.

Мы так же заостряем внимание схеме блокировки IP адресов, с которых CAPTCHA многократно вводилась неверно. Неверные значение CAPTCHA могут вводиться сколько угодно раз, причем как человеком, так и роботом, и это никак не влияет на процесс работы системы. Скажем больше, CAPTCHA как раз на это и рассчитана. Более корректно, по нашему мнению, фильтр должен работать при следующем раскладе: система блокирует так же те IP адреса, с которых за определенный промежуток времени значение CAPTCHA было введено верно множество раз. Иными словами, подозрительно, если пользователь, регистрируется на одном и том же сайте, с одного и того же IP адреса множество раз за короткий промежуток времени. При этом, естественно каждый раз данные капчи вводятся / указываются верно.

Мы так же считаем, что разработчикам безусловно надо учитывать потребности, и что более важно, возможности среднестатистических пользователей. И основываясь на этом конструировать в первую очередь удобный, и эффективный механизм защиты. Удобство потребления является приоритетом для любого проекта, и потеря потенциальных пользователей из-за преград в виде сильного усложнения процесса регистрации, что на сегодняшний день достаточно актуально - является проигрышным вариантом для любого проекта.

Библиографический список:

1. Captcha development problems // Modern technics and technologies. 2015. № 7 [Electronic journal]. URL: <http://technology.snauka.ru/en/2015/07/7577>
2. Designing a Secure Text-based CAPTCHA // Procedia Computer Science. Volume 57, 2015, Pages 122-125