# Введение

CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) или HIP (Human Interactive Proof) - это автоматический механизм безопасности, позволяющий отличить, кто является пользователем - человек или компьютерная программа. Он создает и оценивает тесты, которые могут быть решены человеком, но находятся за пределами возможностей современных компьютерных программ. Он превратился в наиболее распространенную стандартную меру безопасности для предотвращения атак автоматических компьютерных программ. С развитием веб-сервисов атаки типа "отказ в обслуживании" (DoS) со стороны вредоносных автоматических программ стали серьезной проблемой, и тест Тьюринга превратился в важнейший способ отличить человека от опасных автоматических программ. В оригинальном тесте Тьюринга судья-человек уполномочен задать ряд вопросов двум игрокам, один из которых - компьютер, а другой - человек, и отличить их друг от друга. CAPTCHA, как и тест Тьюринга, отличает людей от компьютеров, но судьей теперь является машина. В общем случае CAPTCHA - это криптографический протокол, в основе которого лежит предположение о сложности, основанное на проблеме ИИ. CAPTCHA подразумевает беспроигрышную ситуацию: либо капча не взломана и существует способ отличить человека от компьютера, либо капча взломана, и сложная проблема ИИ решена. CAPTCHA обычно представляет собой простой визуальный тест или головоломку, которую человек может пройти без особого труда, но автоматическая программа не может понять сложный фон, чтобы сделать их трудночитаемыми для программ оптического распознавания символов (OCR). CAPTCHA находит широкое применение в Интернете и других приложениях, таких как защита от червей и спама, онлайн-опросы, бесплатные почтовые сервисы, предотвращение атак на словарь, а также играет важную роль в ограничении скорости использования.

HAI (Human Artificial Intelligence) исследует взаимодействие между людьми и компьютерами, а также основные явления, которые их окружают. Он обозначает характеристики удобства использования, которые тесно связаны с пользовательским интерфейсом и человеческими факторами. Следовательно, она глубоко связана с информатикой, искусственным интеллектом и когнитивной психологией. Основное понятие в HAI - удобство использования. С этой точки зрения такие головоломки, как CAPTCHA, которые люди могут легко решить, но компьютеры находят сложными, являются примером HAI. В этом исследовании мы представили обзорный анализ HAI в аспекте безопасности, а именно: открытые проблемы, трудности и возможности текущих схем CAPTCHA. Остальная часть данной работы организована следующим образом: В разделе II представлена таксономия CAPTCHA-атак. В разделе III описывается анализ проблемы CAPTCHA. В результате в разделе IV приводятся предложения и рекомендации по созданию хорошей CAPTCHA. Наконец, в разделе V приводится заключение.

* 1. **Эволюция CAPTCHA**

Первый человек, Мони Наор, предложил теоретические подходы для отличия компьютеров от людей. В 1997 году поисковая система AltaVista стала первой, кто использовал CAPTCHA в Интернете. Текстовые CAPTCHA были ведущей техникой в начале 2000-х годов. Был разработан ряд атак, использующих обработку изображений, распознавание образов и алгоритмы машинного обучения (ML) для взлома популярных текстовых схем. Кроме того, в попытке повысить безопасность существующих текстовых CAPTCHA использовались алгоритмы антираспознавания и антисегментации. В 2014 году компания Google сообщила, что разработки в области ИИ позволяют распознавать искаженные варианты текста с вероятностью 99,8 %. С 2004 года проблемы компьютерного зрения (КЗ), включая классификацию и распознавание изображений, считались более сложными задачами ИИ, чем распознавание текста. После этого появилось множество схем CAPTCHA на основе изображений с перетаскиванием, выбором изображения или скольжением, чтобы отличить человека от компьютера. Однако передовые решения в области CV и ML помогли победить самые важные схемы CAPTCHA на основе изображений в период с 2013 по 2018 год. Несколько схем CAPTCHA на основе изображений, например схема reCAPTCHA V2, были атакованы спомощью ML. Кроме того, в качестве мер противодействия моделям глубокого обучения были предложены такие подходы, как искажение, смешивание фонового шума и использование неблагоприятных примеров. Для повышения безопасности от атак на основе ML были предложены состязательные примеры, разработанные Сегеди и другими предложили решатель CAPTCHA, который использует инкрементное обучение на ограниченном наборе данных, чтобы победить противные CAPTCHA. Чтобы справиться с пользователями с ослабленным зрением, исследователи предложили CAPTCHA на основе звука в дополнение к CAPTCHA на основе текста и изображений. Однако языковые барьеры и низкое удобство использования ограничивают эффективность этих схем. Кроме того, контролируемое обучение и автоматическое распознавание речи (ASR) показывают, как эти схемы могут быть использованы. Исследователи начали разрабатывать схемы CAPTCHA на основе поведения в 2010-х годах, чтобы создать трудности, основанные на поведенческих особенностях. Первая поведенческая CAPTCHA была запущена компанией Geetest в 2012 году, в то время как Google выпустила No CAPTCHA reCAPTCHA в 2014 году и невидимую CAPTCHA в 2015 и 2017 годах. Было доказано, ч т о атаки ботов, имитирующих поведение пользователя, уязвимы для этих схем. Из-за серьезных проблем с конфиденциальностью Cloudflare недавно решила отказаться от использования reCAPTCHA. Наконец, последние направления исследований используют данные датчиков для создания задач, которые сложно воспроизвести автоматическим ботам. Однако нам необходимо подождать достаточное количество времени, прежде чем мы сможем в полной мере оценить CAPTCHA, основанные на сенсорных данных.

* 1. **Коды CAPTCHA**

Схемы CAPTCHA меняются и постоянно совершенствуются в результате развития передовых технологий, искусственного интеллекта и методов взлома. Основные коды CAPTCHA в настоящее время классифицируются как когнитивные/поведенческие, основанные на видео, аудио, изображениях, тексте и другие.

* + 1. ***Текстовые CAPTCHA***

С годами эти CAPTCHA стали применяться все чаще. В этих методах текст искажается и показывается пользователю в виде изображения, а пользователь должен точно ввести этот текст, чтобы пройти тест. ИИ предполагает, что человек может легко прочитать искаженный текст, но боты, использующие методы оптического распознавания символов (OCR), считают это сложным. Различные визуализации текста задачи можно разделить на три подкатегории: 2D, 3 D и анимация.

*Двухмерные текстовые CAPTCHA* Андрей Бродер и его команда в DEC Systems Research Center изобрели схему двухмерных текстовых CAPTCHA в 1997 году. Похожий метод использовался сайтом AltaVista для предотвращения влияния ботов на рейтинг сайтов в поисковой системе. Фон Ан и Блюм создали Gimpy CAPTCHA и EZ-Gimpy в сотрудничестве с Yahoo в 2000 году, чтобы предотвратить создание ботами вредоносных рекламных объявлений и бесплатных аккаунтов. Gimpy CAPTCHA требует, чтобы вы правильно набрали хотя бы три из семи случайных слов в словаре. EZ-Gimpy - это сокращенная версия Gimpy, показывающая только одно случайное слово из словаря. При создании изображений слов используются различные шрифты, градиенты, шум и другие эффекты, чтобы ботам было сложно их распознать. В 2003 году Моника Чу и Генри Бэрд предложили BaffleText - текстовую CAPTCHA, использующую произносимые псевдослучайные слова с алгоритмами маскировки для предотвращения распознавания программным обеспечением OCR. В 2010 году Megaupload.com создал схему CAPTCHA, устойчивую к сегментации. В этом методе используются перекрывающиеся символы, а также принцип "гештальт-восприятия". Согласно принципу гештальт - восприятия, люди могут полностью восстановить отдельные символы, в то время как компьютеры все еще не справляются с этой задачей. Первая версия ReCAPTCHA была разработана для защиты сайтов от компьютерных атак. Если пользователь правильно набирает известные слова из двух искаженных слов из старых книг, он проходит испытание. Чоу и другие предложили концепцию текстового CAPTCHA с возможностью клика. Их подход предполагает построение сетки кликабельных CAPTCHA из множества текстовых CAPTCHA-задач. Пользователь должен выбрать элементы сетки, которые соответствуют требованиям задачи. Вместо использования машинного текста авторы работ предложили рукописные CAPTCHA для предотвращения распознавания программным обеспечением OCR.

*3D-текстовые CAPTCHA* эти схемы CAPTCHA используют преимущества последовательностей 3D- символов, которые распознаются людьми, но не распознаются ботами, что делает их лучше 2D-текстовых CAPTCHA. Исследовательская группа OCR разработали Teabag3D, высокозащищенную CAPTCHA. Эта CAPTCHA состоит из изображения, смешивающего текстовые символы с трехмерным рисунком. Super CAPTCHA и 3DCAPTCHA - это текстовые схемы CAPTCHA, использующие те же предположения, что и Teabag3D. С 2013 года Super CAPTCHA доступна в виде плагина для WordPress.org. Imsamai и Phimoltares разработали схему 3D CAPTCHA, которая предполагает показ трехмерных буквенно-цифровых последовательностей и смешивание множества эффектов, таких как наложение, поворот, шум, изменение шрифта, масштабирование и другие эффекты, чтобы обмануть распознавание автоматических ботов. Сузи и другие недавно предложили DotCHA, 3D-текстовую CAPTCHA. 3D-буквы состоят из маленьких сфер в каждом задании. Каждая буква читается под разным углом поворота вокруг горизонтальной оси. В результате 3D-текстовые модели необходимо несколько раз повернуть, чтобы определить их буквы.

*Анимированные текстовые CAPTCHA* эти CAPTCHA добавляют временное измерение к текстовым схемам. В частности, текстовое содержимое анимируется в коротком ролике для каждого задания, что усложняет процесс извлечения для ботов с автосоставом. В 2006 году Фишер и Херфет предложили одну из первых анимированных CAPTCHA. Концепция этой CAPTCHA заключается в проецировании текста на анимированную деформирующуюся поверхность. В 2009 году Науманн и другие разработали анимированную CAPTCHA на основе идеи восприятия глазной системы человека. Только когда буквы двигаются, пользователи могут отличить текст от фона. В соответствии с той же концепцией Цуй и другие представили анимированную CAPTCHA, которая распознает правильные символы только при движении. Принцип "ноль знаний на кадр" применяется для того, чтобы исключить утечку информации в каждом кадре. В 2010 году Creo Group выпустил а анимированную HelloCAPTCHA. Для каждой задачи последовательность из шести символов представляется в виде GIF-изображения с некоторыми эффектами: случайными позициями, различными ориентациями и другими. Информация выравнивается и распределяется по нескольким кадрам, чтобы предотвратить распознавание по одному кадру. Задача в NuCaptcha начинается с видео с движущимся текстом с белым шрифтом, за которым следуют три красных символа на динамичном фоне. Чтобы пройти испытание, пользователь должен правильно набрать красные символы. В Dracon CAPTCHAs пять символов отображаются в фиксированных местах, которые случайным образом изменяются с эффектами затухания, размытия и шума.

* + 1. ***CAPTCHA на основе изображений***

В связи с недавним провалом почти текстовых CAPTCHA все большее беспокойство вызывают надежность их защиты и доступность. В последнее время все больше разработок сосредоточено на распознавании изображений, а не символов с предположением, что задачи общего зрения сложнее, чем распознавание текста.

*Интерактивные CAPTCHA* эти CAPTCHA основаны на взаимодействии с пользователем, например жестах пролистывания или движениях мыши, для выявления скрытых точек на изображении. Конти и другие предложили CAPTCHaStar, в которой используется способность человека распознавать фигуры в загроможденной среде. Задача CAPTCHaStar состоит из белых пикселей, называемых звездами, которые случайным образом перемешаны между собой. Положение этих звезд меняется в зависимости от того, где находится курсор. Чтобы пройти тест CAPTCHA, пользователь должен перетащить курсор так, чтобы звезды образовали понятную форму, а затем нажать левую кнопку мыши. Окада и другие создали Noise CAPTCHA с той же концепцией. Эта CAPTCHA состоит из двух частей изображения разного размера с шумом, а также скрытый объект или сообщение на одном из изображений. Пользователи должны перетащить маленькое зашумленное изображение, чтобы определить скрытый объект на большом изображении, прежде чем нажать кнопку "отправить", чтобы пройти испытание CAPTCHA. Cursor CAPTCHA, предложенная Томасом и другими, отображает пять курсоров в случайном порядке на сгенерированном изображении. Чтобы пройти испытание, пользователь должен навести указатель мыши на определенный курсор.

*CAPTCHA, основанные на выборе* эти CAPTCHA требуют от пользователей выбрать изображения-кандидаты из набора изображений. Для описания этой задачи можно использовать только текст или текст с образцом изображения. Asirra - типичная CAPTCHA этой схемы, в которой пользователей просят выбрать всех кошек из набора 12 изображений собак и кошек. В CAPTCHA HumanAuth от пользователей требуется подобрать все изображения, содержащие естественное содержание, среди изображений естественного содержания (например, дерево или река) с изображениями искусственного содержания (например, автомобиль или часы). SEMAGE (SEmanti- cally MAtching imaGEs) CAPTCHA отличается от Asirra и HumanAuth CAPTCHA тем, что требует о т пользователей выбирать семантические изображения из набора изображений. В результате пользователь должен сначала распознать содержание каждого изображения, а затем определить семантическую связь между ними. В 2014 году Google выпустила "No captcha reCAPTCHA". Анализируя окружение браузера (например, cook- ies и историю браузера), система определяет, сталкивается ли она с ботом или нет. В зависимости от степени риска на странице отображается только чекбокс или CAPTCHA на основе выбора. При прохождении CAPTCHA на основе выбора выдается девять изображений-кандидатов и образец изображения с описанием требуемого содержания. Чтобы пройти испытание, пользователь должен выбрать изображения, схожие с образцом. Образ CAPTCHA от Facebook по своему подходу похож на reCAPTCHA. Чтобы пройти испытание, пользователи должны выбрать изображения, соответствующие описанию подсказки, из набора двенадцати изображений с различными контенциями. Avatar CAPTCHA просит пользователей выбрать лица аватаров из набора 12 полутоновых изображений, включающих как человеческие лица, так и лица аватаров. FR-CAPTCHA и FaceD- CAPTCHA - еще две CAPTCHA с изображением лица. FR- CAPTCHA требует от пользователей подобрать два изображения лица одного и того же человека на сложном фоне. С другой стороны, в FaceDCAPTCHA пользователи должны выбрать между визуально искаженными изображениями человеческого лица и нечеловеческого лица.

*CAPTCHA на основе кликов* эти схемы отображают текст и изображение, указывающее, куда должен нажать пользователь, чтобы пройти испытание. Основное ограничение этого типа заключается в том, что для создания нового экземпляра требуется вмешательство человека. Неявные CAPTCHA

является распространенным примером, который требует от пользователей щелкнуть на идентичном месте изображения. Танг и др. впервые использовали SACaptcha, в которой для прохождения испытания пользователи должны нажимать на некоторые области CAPCHA, связывающие объясненную специфическую форму.

*CAPTCHA на основе рисунка* В 2006 году Ширали-Шахреза первым разработал Drawing CAPTCHA, CAPTCHA на основе рисунка. Ромбовидные точки соединяются линиями рисунка пользователя. Наиболее сложным аспектом является то, что пользователи должны идентифицировать эти точки на шумном фоне. VAPTCHA (Variation Analysis-Based Public Turing Test to Tell Computers and Humans Apart) состоит из изображения со случайно сгенерированным траекторией в задании. Чтобы пройти тест, пользователь должен нарисовать траекторию, совпадающую с этой траекторией. В MotionCAPTCHA пользователей также просят нарисовать фигуру, похожую на ту, что изображена в окне с заданием.

В этих CAPTCHA *на основе слайдов*, чтобы решить задачу, пользователи должны использовать слайдер, например, перетащить фрагмент изображения в нужное место, повернуть изображение в нужную ориентацию или выбрать правильную форму изображения. WHAT's Up CAPTCHA показывает три повернутых изображения в случайном порядке, и пользователи должны повернуть их в правильное положение. CAPTCHA Minteye's Slide-to-Fit отображает закрученное изображение, и пользователи должны перемещать ползунок, пока не увидят неискаженную версию изображения. Tencent CAPTCHA требует от пользователей передвинуть ползунок, чтобы совместить две части головоломки. В этих CAP*TCHA* от пользователей требуется совместить части изображения, чтобы сформировать целостное изображение, перетаскивая их. Garb CAPTCHA отображает четыре случайно перемешанные части изображения. Пользователям необходимо изменить порядок расположения этих частей изображения, чтобы получить полное изображение для прохождения теста CAPTCHA. Хамид Али и другие впервые использовали CAPTCHA на основе головоломки. Для выполнения задания необходимо перетащить четыре фрагмента изображения в пустую сетку из четырех ячеек. Гао и др. предложили CAPTCHA на основе изображения в виде пазла. В этой CAPTCHA изображение делится на множество частей (например, 9, 16 или 25), и только две части располагаются неправильно. Пользователям необходимо поменять местами неправильные части, чтобы решить задачу. Capy CAPTCHA требует от пользователей переместить кусочек головоломки в недостающее место в задаче. Пропавшее место заполняется случайным изображением. KeyCAPTCHA показывает три кусочка головоломки и неполное изображение. Пользователю необходимо собрать эти кусочки так, чтобы они соответствовали эталонному изображению. Как только курсор останется в кадре, эталонное изображение исчезнет. Чтобы пройти испытание CAPTCHA, пользователь должен переставить эти кусочки в правильные места.

* + 1. ***Аудиокапчи наоснове CAPTCHA***

Для людей с нарушениями зрения в качестве альтернативы визуальным схемам CAPTCHA б ы л и предложены схемы CAPTCHA на основе звука. Чтобы пройти тест, они должны набрать услышанное. В Университете Карнеги-Меллона исследователи представили звуковую reCAPTCHA, которую позже приобрела компания Google. Чтобы решить задачу, пользователи должны определить восемь цифр, произнесенных в человеческом шуме, и принять только одну неправильную цифру из этих цифр. Аудиокапча eBay состоит из шести цифр, произнесенных разными шумными голосами. Microsoft CAPTCHA состоит из десяти цифр, произнесенных разными голосами, смешивающимися с шумом некоторых разговоров. Yahoo CAPTCHA требует о т пользователей ввести семь цифр после трех детских звуковых сигналов с фоновым шумом. Версия Audio reCAPTCHA 2013 года требует от пользователей распознать все цифры, разделенные на три кластера в задании. В каждом кластере можно найти три или четыре совпадающие цифры. Новая версия reCAPTCHA 2017 года включает десять произносимых цифр и фоновый шум.

* + 1. ***CAPTCHA на основе видео***

В задаче создается короткий видеоролик, отражающий определенное содержание, от пользователей требуется понять и описать его текстом. Клювер и другие предложили CAPTCHA, в которой с помощью короткого видео пользователи должны посмотреть и затем ввести три слова для его описания. Ширали-Шахреза и др. предложили CAPTCHA движения. которая требует от пользователей описать движение человека в просмотренном видео, выбрав одно из предложений.

* + 1. ***Когнитивные CAPTCHA***

Методы CAPTCHA, основанные на когнитивных способностях и обеспечивающие повышенную безопасность, в значительной степени заменили традиционные методы Captcha. Когнитивные способности - это навыки, основанные на работе мозга, которые являются результатом определенного сочетания нейробиологических и психологических методов. Знания, концентрация, память, суждения и оценки, рассуждения и вычисления, решение проблем и принятие решений - все это аспекты человеческого познания и поведения. Чтобы отличить человека от бота, эти методы CAPTCHA используют биометрические (то, кем вы являетесь), физические (то, что у вас есть) и основанные на знаниях (то, что вы знаете) факторы с поддержкой или без поддержки таких датчиков, как гироскоп или акселерометр. В 2020 году Асьен и другие предложили BeCAPTCHA-Mouse, которая отличает людей от ботов, анализируя траектории движения мыши во время выполнения задания. Gametrics отличает людей от ботов, собирая и анализируя движения мыши пользователя во время операций перетаскивания для решения динамической когнитивной игры. GEETest и Netease. как и Tencent CAPTCHA, требуют от пользователей пройти CAPTCHA на основе скользящего изображения, перемещая ползунок до тех пор, пока не совпадут два фрагмента головоломки. Если пользователи выполняют задание и их поведение при скольжении не вызывает подозрений, считается, что они прошли испытание. Сирипитак-чай и другие предложили EYE- CAPTCHA, в которой пользователи должны решить математическую CAPTCHA, двигая глазами. Чтобы решить задачу, пользователь должен определить правильный ответ и с помощью глаз переместить его в центр экрана. В 2014 году Google запустил функцию "No CAPTCHA reCAPTCHA" (reCAPTCHA V2). Все, что требуется, - это поставить галочку в поле "Я не робот", однако поведение пользователя (нажатия, перемещения мыши и другие действия), а также другая информация (браузер, куки, история и т. д.) собираются и анализируются в фоновом режиме. Если есть подозрение, что пользователь является ботом, ему необходимо пройти вторую проверку reCAPTCHA, основанную на изображении. В 2017 году была выпущена Invisible reCAPTCHA, обновленная версия reCAPTCHA V2. Процесс оценки запускается в фоновом режиме с помощью вызова JavaScript API или нажатия пользователем на существующую кнопку. Невидимая reCAPTCHA, как и подход "No CAPTCHA reCAPTCHA", требует второго задания reCAPTCHA на основе изображения, если есть подозрение, что пользователи являются ботами. В 2015 году Герар и другие первыми представили физическую CAPTCHA для мобильных устройств, названную CAPPCHA (Completely Automated Pub- lic Physical test to tell Computers and Humans Apart). Пользователи должны наклонить устройство на определенный градус, что сложно сделать ботам. Хапперих и др. В 2016 году была представлена Sensor CAPTCHA, в которой от пользователей требуется выполнить сложный жест (например, ловить рыбу, б и т ь молотком, пить) с помощью мобильного устройства. Авторы работы предложили Pedometric CAPTCHA, в которой о т человека требуется пройти не менее пяти шагов. Когда пользователь ходит, в мобильном устройстве возникает ускорение, что затрудняет работу ботов. Мантри и другие предложили схему CAPTCHA, в которой пользователи должны выполнить требование о перемещении устройства в соответствии с определенным руководством, отображаемым на устройстве. Франк и другие проинструктировали пользователей выполнить определяемый жест, регистрируемый гироскопом (например, вращение, наклон, или рисунка), при перемещении устройства. Герар и другие разработали Invisible CAPPCHA, которая похожа на CAPP- CHA т е м. что задача невидима для пользователей. Датчики считывания обнаруживают касания пользователя, а не события на сенсорном экране, которые боты могут легко имитировать. Кроме того, эта CAPTCHA защищает конфиденциальность пользователя, не отправляя конфиденциальные данные на сервер. AccCAPTCHA требует от пользователя сыграть в игру с катящимся шариком. Чтобы завершить игру, пользователь должен управлять мячом с помощью датчиков движения устройства. GISCHA, CAPTCHA на основе игры для мобильных устройств, была предложена Янг и др. Чтобы пройти испытание, пользователь должен переместить мяч в правильное отверстие. Абабтейн и другие предложили CAPTCHA, которая требует от пользователей пройти простую игру с использованием сенсоров. Они предложили пять игр, в каждой из которых есть несколько статичных и один движущийся объект. Пользователи должны перемещать движущийся объект, чтобы попасть в правильные статические объекты, чтобы пройти испытание. SenCAPTCHA была предложена Fengetal. для определения местоположения ключевой точки на лице животного. Пользователям показывают небольшой красный шарик и изображение животного. Затем они должны направить красный шарик в центр глаза животного, наклоняя свои устройства. Авторы предложили BrightPass, мобильный CAPTCHA для защиты PIN/пароля при аутентификации. Предложенный ими механизм использует яркость экрана, которую автоматические боты не могут обнаружить, чтобы определить, когда пользователь должен ввести правильную цифру или обманчивую. В виде физической CAPTCHA авторы предложили CAPTCHA для аутентификации на основе PIN-кода, используемую для смарт-часов. Этот механизм основан на той же концепции, что и CAPPCHA. Для ввода пароля н е о б х о д и м о физически повернуть безель на определенный градус. Аналогичным образом авторы используют поворот цифровой заводной головки в смарт-часах для защиты PIN-кода.

* + 1. ***Другие типы***

Стефан Поповеньюк в 2010 году предложил метод аутентификации SpeakUP для удаленного голосования без наблюдения. Голосовая биометрия дополнена текстовой CAPTCHA. Избиратели должны вслух прочитать характеристики проголосовавшего кандидата, представленные в виде двумерной текстовой CAPTCHA. Кроме того, голосовые биометрические характеристики избирателя определяются с помощью задачи. Автор также предложил записывать видео решения задач избирателем. Для защиты систем аутентификации по лицу Узун и другие предложили rtCaptcha - Real-Time CAPTCHA. Пользователи должны записать свое произношение вслух представленной двумерной текстовой CAPTCHA.

# Анализ атак CAPTCHA

CAPTCHA превратилась в наиболее популярную стандартную меру безопасности для предотвращения автоматизированных атак компьютерных программ. В последние годы многие методы атак, разработанные хакерами или исследователями, позволили эффективно взломать все распространенные традиционные схемы. Некоторые методы, в том числе Invisible reCAPTCHA, до сих пор не взломаны. Однако с появлением ботов четвертого поколения, точно имитирующих поведение человека, безопасная CAPTCHA вряд ли будет создана без дополнительных специальных устройств. В частности, почти все когнитивные CAPTCHA с сенсорной поддержкой пока не уязвимы для автоматизированных атак. Тем не менее, они все еще подвержены атакам ретрансляции с помощью человека, поскольку имеют ограниченное количество задач и могут быть решены только с помощью доверенных устройств.

* 1. **Атака на текстовые CAPTCHA**

Текстовые CAPTCHA были первой схемой CAPTCHA и до сих пор остаются самыми популярными. Мори и Малик в 2003 году в работе был представлен метод атаки на соответствие формы, который позволил пройти капчи Gimpy и EZ-Gimpy с точностью 33% и 92% соответственно. В предложенном методе использовались алгоритм корреляции и алгоритм прямой оценки искажений для успешного взлома EZ- Gimpy с вероятностью 99 %. Чел- лапилла и другие создали высокозащищенную CAPTCHA анти- сегментации в 2005 году после прохождения различных текстовых CAPTCHA с помощью машинного обучения. В 2008 году Эль Ахмад и Ян продемонстрировали возможность взлома нескольких антисегментационных CAPTCHA, используемых Google, Microsoft и Yahoo. Позже другие исследователи пытались пройти эти CAPTCHA с более высокими показателями успеха. Эль Ахмад и Ян также взломали CAPTCHA Megaupload с 78 % успеха. Исследователи Google использовали нейронные сети для преодоления самой сложной категории ReCAPTCHA в 2014 году с точностью 99,8 %. Авторы предложили методы атаки 3D CAPTCHA без использования программного обеспечения OCR. В нескольких 3D CAPTCHA, таких как 3DCAPTCHA, Teabag 3D и Super CAPTCHA, они извлекали пиксели из символов для автоматического распознавания задачи. Используя эту технику, авторы смогли взломать 3DCAPTCHA, Teabag 3D и Super CAPTCHA с коэффициентами успешности 58 %, 31 % и 27 % соответственно. Более того, те же авторы смогли пройти Teabag 3D, используя информацию о боковой поверхности 3D-текстовых объектов. В случае с анимированными CAPTCHA Нгуен и другие продемонстрировали, как легко извлечь информацию из нескольких анимированных кадров с помощью CL (Catch- ing Line) или PDM (Pixel Delay Map). Эти методы успешно справились с такими анимированными CAPTCHA, как KillBot Professional, iCAPTCHA, Dracon CAPTCHA и Atlantis. Из-за уязвимости к атакам сегментации эти же методы были использованы для поражения вариантов HelloCAPTCHA с успешностью от 16 до 100%. NuCaptcha - это анимированная CAPTCHA, устойчивая к сегментации, которая работает путем наложения и сцепления друг с другом, чтобы противостоять методам атаки PDM или CL. Эли Бурштейн разделил объекты в каждом кадре с вероятностью успеха 90 %, используя оценку плотности точек интереса (алгоритм SIFT) и анализ формы ограничительной рамки.

* 1. **Атака на CAPTCHA на основе изображений**

Голле удалось взломать схему Асирры. Для этого была использована SVM (машина опорных векторов) для классификации кошек и собак с успешностью 82,7 %. Эрнандес-Кастро и другие предложили атаку по побочным каналам для взлома HumanAuth с точностью 92 %. CAPTCHA на основе изображений Facebook и CAPTCHA на основе изображений Google были обойдены Sivakorn et al. с показателями успешности 83,5 % и 70,78 % соответственно. Авторы добились успеха в 79 и 88 % с новой и старой вариациями reCAPTCHA V2. Они также победили China Railway CAPTCHA и Facebook image CAPTCHA с показателями успешности 90 и 86 % соответственно. Кроме того, эти авторы разрушили различные схемы CAPTCHA на основе изображений, включая Tencent CAPTCHA с успешностью 100 %. Конволюционные нейронные сети (CNN) были применены для успешного взлома Avatar CAPTCHA с коэффициентом успешности 99 %. Как FaceDCAPTCHA, так и FR-CAPTCHA были побеждены Gao et al. с коэффициентами успешности 48 % и 23 % соответственно. Атака Minteye CAPTCHA была побеждена с помощью использования длины краев изображения и операторов Собеля. Метод атаки выбирает изображение с наименьшей суммой краев, основываясь на том, что в закрученном изображении края длиннее. Эрнандес-Кастро и другие предложили недорогую атаку с использованием JPEG для измерения непрерывности изображения. Используя эту атаку по побочным каналам, они успешно взломали Capy CAPTCHA, Garb CAPTCHA и KeyCAPTCHA с коэффициентами успешности 65,1%, 98,1 % и 20 % соответственно. Gougeon и Lacharme недавно смогли победить CAPTCHAaStar с коэффициентом успешности 96 %. Они также продемонстрировали, что настройка параметров не предотвращает атаку этой CAPTCHA на концентрацию пикселей (звезд) во время формирования изображения.

* 1. **Атака на CAPTCHA, основанные на аудиозаписях**

Там и другие экспериментировали с подходом на основе SVM, чтобы победить аудио reCAPTCHA, и получили 45 % успеха для решения с точным совпадением и 58 % успеха для условия "одна ошибка". Decaptcha от Burzstein и Bethard продемонстрировала 75-процентный успех в обходе аудио CAPTCHA от eBay. Их метод анализирует волновой файл с помощью дискретного преобразования Фурье (DFT), а затем группирует энергетические пики. Затем, чтобы распознать речевые паттерны, используется алгоритм контролируемого обучения для обучения аудиоданных. Авторы представили разгадыватель CAPTCHA с непрерывной речью, который разрушил аудио CAPTCHA Yahoo и Microsoft с коэффициентами успешности 45 % и 49 % соответственно. Этап классификации в этом решателе был контролируемым, в то время как этап автоматической сегментации был неконтролируемым. Для разметки использовался Amazon Mechanical Turk, а отсканированные CAPTCHA были классифицированы с помощью алгоритма регуляризованной классификации по методу наименьших квадратов (RLSC). Из-за наличия семантического голосового шума их система смогла решить reCAPTCHA только с коэффициентом успешности 1,5 %. Сано и другие предложили взломщик CAPTCHA для непрерывной речи, чтобы победить антисегментационные CAPTCHA, которые перекрывают целевые голоса. Для распознавания речи использовались скрытые марковские модели (HMM), которые были протестированы на версии аудио reCAPTCHA 2013 года с успешностью 52 %. Бок и другие представили unCaptcha, которая может обойти версию аудио reCAPTCHA 2017 года с коэффициентом успешности 85,15 %, используя бесплатные онлайн-сервисы преобразования речи в текст и выполняя минимальное фонетическое отображение для повышения точности.

* 1. **Атака на когнитивные CAPTCHA**

Используя четыре функции моделирования (Softmax, Sigmoid, Tanh и ReLu) для имитации человеческого поведения, Чжао и др. успешно обошли скользящие CAPTCHA, такие как GeeTest и Netease CAPTCHA, с показателями успешности 96 и 98 % соответственно. Создав куки для отслеживания автоматических ботов, Сивакорн и другие смогли обмануть систему анализа рисков Google. В результате после 9 дней просмотра автоматическими ботами различных сервисов Google решатель может поставить галочку в поле "Я не робот". Кроме того, авторы предложили простую атаку с показателем успешности 70,78 % для преодоления второй задачи reCAPTCHA V2. Чтобы взломать No CAPTCHA reCAPTCHA, необходимо авторы применили стратегию "разделяй и властвуй". Они добились успеха в 97,4 % случаев на сетке 100×100 и в 96,7 % случаев на экране с разрешением 1000 × 1000.

* 1. **Атака на другие CAPTCHA**

Клювер и другие разработали подход, основанный на частоте тегов, для атаки на предложенную ими CAPTCHA на основе видео с успешностью 13 %. Эрнандес-Кастро и другие успешно взломали QRBGS CAPTCHA с помощью атаки по боковому каналу с коэффициентом успеха 44,54 %. Мохамед и другие продемонстрировали, что атаки на основе словарей способны победить DCG CAPTCHA. Более того, разработчики предложили решатель, который автоматически обходит SweetCAPTCHA, различные слайдерные CAPTCHA (схема Taobao) путем разработки простого JavaScript-кода и кукловода.

* 1. **Другие нападения**
     1. ***Атака по побочному каналу***

Атаки по побочным каналам - это процессы, которые пытаются решить проблему, значительно более простую, чем исходная. Предполагаемое решение строится вокруг сложной задачи (AI-hard problem), в то время как фактическое решение строится вокруг любых проблем проектирования или реализации, чтобы избежать более сложного подхода. Для выявления связи между задачами и ответами на них эти атаки опираются на случайные отклонения, отсутствующие в равномерной случайности. В этом случае вызов предоставляет (непреднамеренно, "утечка" или "побочный канал") знания о б ответе. Атаки по побочным каналам в ASIRRA кратко описаны в этом разделе. ASIRRA состоит из более чем 25 000 фотографий, половина из которых классифицируется как кошки или собаки. Эти фотографии были обработаны классификатором, который, не используя никаких методов распознавания образов, смог различить кошачьи и собачьи фотографии с точностью около 60 %. Авторы HumanAuth решили смешать PNG-изображение со случайным JPG-изображением, выбранным из библиотеки, чтобы предотвратить легкое индексирование библиотеки изображений. Выбор нового водяного знака, оказывающего большее влияние на исходное изображение, был бы сделан в ущерб удобству использования.

* + 1. ***Атака на основе признаков***

В 2009 году Филипп Голле представил эффективные атаки на ASIRRA, основанные на анализе особенностей CAPTCHA, таких как шрифт, форма, текстура и цвет. Используя обработку изображений, этот подход делит фотографии на сетку ячеек с текстурой и цветом (серая шкала), которая затем подается на классификаторы опорно-векторной машины (SVM) с успехом классификации 83 %.

* + 1. ***Атака на основе базы данных***

Если CAPTCHA основана на общедоступной базе данных (например, фотографии с метками), существует множество потенциальных атак на эту базу:

* + - * Атаки на индексирование базы данных: база данных tfle может быть загружена (по крайней мере, частично) для получения информации, необходимой для решения tfle CAPTCHA.
      * Атаки, отравляющие базу данных: при открытой и незащищенной базе данных CAPTCHA наша информация может быть загружена, чтобы помочь нам решить CAPTCHA с помощью этой информации.
    1. ***Решающая атака человека***

CAPTCHA предназначены для заполнения людьми, но существуют рынки услуг по решению CAPTCHA.

(обычно в регионах с дешевой рабочей силой) и ретрансляционные атаки, которые передают задачи

CAPTCHA людям, получающим выгоду от их решения.

# Анализ проблемы CAPTCHA

* 1. **Угрозы нападения**

С развитием автоматизированных атак различия в решении CAPTCHA между людьми и автоматизированными ботами могут стать неактуальными: Должен ли человек, просматривающий другой веб-сайт или находящийся в графическом интерфейсе другой программы, не иметь права решать наши CAPTCHA? Является ли компьютерная программа, которой помог человек, все еще автоматической атакой? Поскольку отличить человека от бота сложно, схемы CAPTCHA требуют дополнительных механизмов для повышения их безопасности:

* Измерьте качество, способность или поведение "флумана", чтобы различать флуманов и компьютеры.
* Различайте алгоритмы flumans и fluman-assisted, чтобы предотвратить увеличивающие или fluman-assisted

атаки.

* Предотвращайте атаки ретрансляции, различая flumans wflo see tfle CAPTCHA на tfle оригинальном сайте

CAPTCHA и tflose wflo see it on anotfler site/ interface.

* Предотвращайте атаки на фермы fluman, используя методы, позволяющие предотвратить или затруднить использование ферм решателей при решении tfle CAPTCHA.
  1. **Твердость AI не передается**

Большинство CAPTCHA были уязвимы в результате одной из следующих проблем:

* + 1. В их основе лежит гораздо более конкретная и слабая проблема, чем предполагалось изначально.
    2. Недостатки разработки или реализации делают tflem слизистую более легкой для обхода с помощью процедур, анализирующих tfleir cflallallenges. В результате эти процедуры известны как атаки по боковому каналу, поскольку они пытаются решить гораздо более простую задачу, чем та, которая была задумана разработчиками CAPTCHA.
    3. Сложность нерешенной проблемы ИИ трудно передать в дизайне CAPTCHA. Мы не знаем, как классифицировать или глубоко понять сложность ИИ, поэтому CAPTCHA с вызовом такой сложности ИИ может оказаться недостаточно сложной для автоматических ботов.
  1. **Недостатки конструкции**
     1. ***Необъективное распределение ответов***

Одна из распространенных ошибок - выбор неравномерно распределенного подмножества возможных ответов. Одним из таких примеров является QRBGS (MathCAPTCHA), разработчики которого использовали в арифметических операциях однозначные цифры. В результате ответы, скорее всего, будут небольшими целыми числами. Еще один пример - Megaupload CAPTCHA, которая избегает использования значений O, I, J и 0. Хуже того, она всегда использует схему "три буквы - затем цифра", что делает ее более удобной для пользователя, но в то же время значительно снижает ее эффективность. В задачах Teabag используются только трехсимвольные длины и избегаются символы, которые трудно различить в 3D-проекции. Символы "S", "Z", "3", "P", "b", "w", "M", "t" и "d" встречались более чем в 3 % случаев в выборке из 100 задач, в то время как основной набор из 34 других символов, включая "1" и "0", не появлялся (возможно, чтобы избежать совпадения с "I" и "O").

* + 1. ***Необъективное распределение задач***

Любая предвзятая идея в дизайне CAPTCHA, не основанная на случайности, может позволить проанализировать задачу, что приведет к атакам через побочные каналы или анализу категоризации задач. Поскольку распределение размеров букв в Teabag неравномерно, фронтальные границы символов могут быть выбраны на основе размера их области. Кроме того, существует корреляция пикселей, которая позволяет обнаружить заднюю границу. Простые алгоритмы, такие как непрерывность пикселей, позволяют обнаружить растущие фоновые области. В некоторых случаях часть изображения, не относящаяся к символам, может быть удалена полностью или почти полностью. Другой пример - капча Megaupload, в которой буквы и цифры всегда напечатаны одним и тем же шрифтом - Antique Olive (по определению Identifont). Символы поворачиваются под определенным углом, по часовой стрелке или против часовой стрелки, причем первая буква - по часовой стрелке, а вторая - против часовой. Это также предотвращает наложение более двух символов.

* + 1. ***Соотношение между задачей и ответом***

Задача может предоставлять (непреднамеренно, "утечка" или "побочный канал") информацию, основанную на содержании ответа. Атаки по боковому каналу могут быть использованы для обхода задач путем использования утечки информации.

* + 1. ***Оценка ответа***

Не всегда нужно упрощать CAPTCHA, чтобы определить, верны ли ответы. По возможности избегайте знания того, является ли ответ на задачу правильным или неправильным, или любого другого способа узнать, близок ли он к правильному. Мы можем передавать эту информацию пользователю через промежуточный механизм связи (например, через учетные записи электронной почты, которые также должны контролироваться, чтобы ограничить время отправки сообщений) или передавать ее так, чтобы ее было трудно отличить от автоматических ботов.

* + 1. ***Зависимость от пользователя***

Вообще, делать CAPTCHA зависимой от пользователя - плохая идея, и еще хуже, если эту зависимость можно узнать или угадать. Например, ASIRRA отображает в Petfinder домашних животных, которые находятся недалеко от места нахождения претендента, чтобы увеличить шансы на усыновление питомцев, отображаемых в CAPTCHA (с помощью IP-геолокации). Этот недостаток является критическим, поскольку он способствует многим типам атак, включая отравление базы данных и индексирование базы данных.

* 1. **Недостатки реализации**

Некоторые системы CAPTCHA можно полностью обойти, воспользовавшись идентификатором сессии ранее использованной CAPTCHA. Это связано с плохой реализацией, но еще несколько лет назад это не было редкостью. Некоторые разработчики до сих пор кодируют ответ на задачу в URL или поле формы. Используя эту ошибку, можно запросить множество вызовов с одним и тем же ответом. В результате можно провести среднюю атаку. вычислив медианные значения этих вызовов. Еще одна ошибка в реализации - отправка клиенту хэша ответа, например хэша MD5, в качестве ключа. Если количество ответов ограничено или распределено неравномерно, хэши этих ответов можно легко узнать, чтобы решить задачи. Кроме того, использование небольших фиксированных пулов задач - один из распространенных недостатков реализации. Например, HumanAuth использует менее сотни изображений, даже маскируя их логотипами, которые легко охарактеризовать или проиндексировать. Кроме того, HumanAuth генерирует ответы на вызовы только со значением 0 или небольшим целым числом. Это позволяет осуществить еще один тип атаки: если ответ 0 окажется неудачным, мы ответим серией целых чисел, начиная с наименьших по абсолютной величине. Другая распространенная ошибка заключается в том, что вызовы QRBGS, как пример, не создаются по запросу, а повторяются. Кроме того, некоторые системы используют крайне рискованный метод связи с сервером CAPTCHA, который легко эксплуатировать.

* 1. **Сохранение конфиденциальности пользователей**

В отличие от традиционных схем CAPTCHA, новые сенсорные и поведенческие схемы CAPTCHA, как б ы л о показано, вызывают проблемы с конфиденциальностью, такие как данные о поведении пользователя, файлы cookie и данные сенсоров, отправленные на удаленные серверы. Некоторые исследователи предложили отправлять на сервер только результаты тестов, а не данные датчиков,

в качестве решения. Однако для предотвращения взлома на стороне клиента требуется доверенное оборудование. Таким образом, на этапе разработки новых схем CAPTCHA необходимо тщательно учитывать конфиденциальность пользователей.

* 1. **Совместимость устройств**

Очевидно, что надежная и удобная CAPTCHA должна быть совместима с широким спектром устройств. Наиболее перспективные схемы CAPTCHA, с другой стороны, в значительной степени зависят от одного устройства. Например, схемы CAPTCHA, основанные на динамике прикосновений и касаний или динамике мыши, требуют специализации устройства. Сенсорные схемы CAPTCHA, для которых требуются датчики, встречающиеся только в смарт-часах, планшетах или смартфонах, трудно реализовать на большинстве устройств пользователей.

# Как разработать хороший CAPTCHA

* 1. **Хорошие свойства**

Любой новый дизайн CAPTCHA должен быть запущен в производство на тестовом сайте, без других средств защиты (чтобы сосредоточиться на твердости CAPTCHA), в течение достаточно длительного периода времени, чтобы можно было провести исследование. Новые CAPTCHA должны включать следующие функции для повышения безопасности против автоматических ботов:

* + 1. Во всех параметрах должна быть случайность и равномерное распределение. Например, для текстовой CAPTCHA: равномерное количество областей, линий, пикселей со случайными свойствами (цвет, группа, размер группы и т. д.), переменное количество символов, различные шрифты, размер изображения и т. д.
    2. Не должно быть более простых CAPTCHA-задач: подтипы или альтернативы должны иметь одинаковый уровень сложности (как, например, визуальные и звуковые CAPTCHA).
    3. Задача должна быть максимально приближена к оригинальной проблеме ИИ.
    4. В конструкцию должны быть включены функции, позволяющие обнаружить автоматический обход или предотвратить релейные атаки.
    5. Задачи распределяются равномерно и не зависят от пользователей и ответов. Кроме того, ответы на вопросы распределяются случайным образом и равномерно. Между заданиями и ответами не должно быть никаких статистических связей.
    6. Затрудните автоматическим ботам определение корректности ответов, используя образцы противника, мехфланизмы ответа или методы взаимодействия с серверами CAPTCHA.
  1. **Обеспечение безопасности**
     1. Повторение ответов: если атакующий способен собрать конечное количество заданий с одинаковыми ответами, то можно быть уверенным, что атакующий не сможет создать лучший ответ, чем случайный. Это означает, что нет лучшей атаки, чем метод проб и ошибок.
     2. Повторение Cflallenge: Если в нашей CAPTCHA есть только конечный набор различных задач, и мы не знаем, как и х решать, то не будет лучшей стратегии, чем метод проб и ошибок, с низким процентом успешности.
     3. Отсутствие категоризации: Если наша CAPTCHA будет состоять из различных типов заданий, то не будет никакого способа автоматически различать их или классифицировать сложность различных заданий.
  2. **Тест на безопасность**

Для этого теста предлагается создать достаточно большой набор элементов (T = тест, A = ответ) тестов. Мы ищем неравномерности в этом распределении, используя инструменты общей случайности и статистического анализа:

* Несоответствия в распределении A (потенциальная атака вслепую).
* Несоответствия в распределении T (категоризация типов и анализ Cflallenge).
* Корреляции между T и A (потенциальная атака по боковому каналу).

Эти тесты могут быть выполнены для некоторых простых свойств T, таких как цветовые гистограммы, размеры областей, гистограммы, расстояния между похожими областями, максимум и минимум для блока байт и битовая корреляция с заданными векторами. Это можно использовать для оценки параметров безопасности любого предложения CAPTCHA, избегая таких "подводных камней", как нерелевантные значения параметров, которые приводят к утечке информации.

# Заключение

CAPTCHA - это соревнование между людьми и компьютерами. Компьютеры пытаются имитировать все, что может делать человек. Люди же, напротив, полагаются на твердость ИИ и его способность к познанию, чтобы бросить вызов компьютерам. Очевидно, что с быстрым и непрерывным развитием технологий компьютеры, оснащенные самым надежным и передовым программным и аппаратным обеспечением, способны в любой момент решить самые сложные проблемы ИИ. В этой статье мы представили обзорный анализ взаимодействия ИИ между пользователями и компьютерами в рамках существующих схем CAPTCHA' аспекты безопасности, открытые проблемы, трудности и возможности в разработке CAPTCHA. Мы ожидаем, что эта работа послужит хорошей отправной точкой для новых разработчиков CAPTCHA, чтобы избежать некоторых общих недостатков проектирования и реализации, а также для разработки новых методологий оценки безопасности и уровня уверенности.