Лежанков М.Е. Исследование эффективности HONEYPOT-решений в ВЕБ-приложениях // Международный журнал информационных технологий и энергоэффективности. – 2026. – Т. 11 № 6(68) Ч.1 с. 45–56

Международный журнал информационных технологий и энергоэффективности

Сайт журнала:
http://www.openaccessscience.ru/index.php/ijcse/



УДК 004.056.57:004.77

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ HONEYPOT-РЕШЕНИЙ В ВЕБ-ПРИЛОЖЕНИЯХ

Лежанков М.Е.
ФГАОУ ВО "НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО" , Санкт-Петербург, Россия (197101, город Санкт-Петербург, Кронверкский пр-кт, д. 49 литер а), e-mail: lemmaksim44@gmail.com

В данной статье представлено экспериментальное исследование эффективности honeypot-решений для защиты веб-приложений от автоматизированных угроз. Проведен анализ отечественных и зарубежных научных работ, посвященных применению honeypot-систем, CAPTCHA и методов обнаружения веб-ботов. В рамках исследования разработан тестовый стенд с различными типами веб-форм, включающими классические honeypot-ловушки, Google reCAPTCHA v2 и Cloudflare Turnstile. Для проверки эффективности защитных механизмов использовались простые HTTP-боты и более продвинутые headless-боты, имитирующие поведение реального пользователя. На основе серии экспериментов выполнен сравнительный анализ решений по критериям скорости обнаружения, устойчивости, уровня ложных срабатываний, производительности и качества собираемых данных. Результаты показали, что классические honeypot-механизмы эффективны преимущественно против простых ботов, однако теряют эффективность при использовании современных средств автоматизации. Наиболее высокие показатели защиты продемонстрировали CAPTCHA-системы, особенно Cloudflare Turnstile, обеспечивающая высокий уровень обнаружения автоматизированных запросов при минимальном влиянии на пользовательский опыт.

Ключевые слова: Honeypot, веб-приложение, информационная безопасность, CAPTCHA, веб-боты, headless-боты, Cloudflare Turnstile, Google reCAPTCHA, защита веб-форм, автоматизированные атаки.

INVESTIGATION OF THE EFFECTIVENESS OF HONEYPOT SOLUTIONS IN WEB APPLICATIONS

Lezhankov M.E.
"ITMO NATIONAL RESEARCH UNIVERSITY", St. Petersburg, Russia (197101, St. Petersburg, Kronverksky prospekt, 49 litera a), e-mail: lemmaksim44@gmail.com

This article presents an experimental study of the effectiveness of honeypot solutions for protecting web applications from automated threats. The analysis of domestic and foreign scientific papers devoted to the use of honeypot systems, CAPTCHA and web bot detection methods is carried out. As part of the research, a testbed has been developed with various types of web forms, including classic honeypot traps, Google reCAPTCHA v2 and Cloudflare Turnstile. To test the effectiveness of the protection mechanisms, simple HTTP bots and more advanced headless bots that mimic the behavior of a real user were used. Based on a series of experiments, a comparative analysis of solutions was performed according to the criteria of detection speed, stability, false alarm rate, performance and quality of the data collected. The results showed that classic honeypot mechanisms are effective mainly against simple bots, but they lose effectiveness when using modern automation tools. The highest protection rates were demonstrated by CAPTCHA systems, especially Cloudflare Turnstile, which provides a high level of detection of automated queries with minimal impact on the user experience.

Keywords: Honeypot, web application, information security, CAPTCHA, web bots, headless bots, Cloudflare Turnstile, Google reCAPTCHA, web form protection, automated attacks.

В условиях цифровизации и стремительного развития информационных технологий наличие корпоративного веб-сайта стало обязательным элементом деятельности большинства современных компаний. Такие сайты могут выполнять как информационные функции в виде лендинг-страниц, так и представлять собой сложные веб-приложения, включая интернет-магазины, сервисные платформы и т.п. Компании активно стремятся перенести процессы и данные в онлайн-среду, обеспечивая удобный доступ к информации для клиентов и партнеров.

Однако открытость веб-ресурсов делает их уязвимыми. Для противодействия угрозам разработчики внедряют различные средства защиты. Одним из эффективных решений являются honeypot-системы – элементы пассивной защиты, встроенные в архитектуру веб-приложения. Они могут использоваться как для выявления несанкционированных попыток доступа к административным панелям и API, так и для защиты от атак, связанных с загрузкой вредоносных файлов, а также для предотвращения эксплуатации известных уязвимостей.

Для проведения исследования необходимо провести анализ отечественных научных источников, который охватывает большое количество тем связанных с системами Honeypot в веб-приложениях. Одним из важных пунктов является оценка актуальности и перспективности технологии. Работа «Анализ ханипотов с открытым исходным кодом» [1] за авторством А.А. Сергановского и В.Н. Романовича содержит анализ репозиториев платформы GitHub по запросу «honeypot». В работе представлены графики, отражающие результаты проведенного анализа, рассмотрим один из них. На рисунке 1 представлена гистограмма, демонстрирующая количество всех проанализированных, уникальных honeypot. Как видно из графика, веб-ловушки, направленные на имитацию веб-приложений и сайтов, занимают одну из лидирующих позиций. В исследовании авторы упоминают, что за последнее время увеличилась активность обновлений репозиториев. Таким образом еще раз доказывается актуальность исследований honeypot-систем в области информационной безопасности веб-приложений.

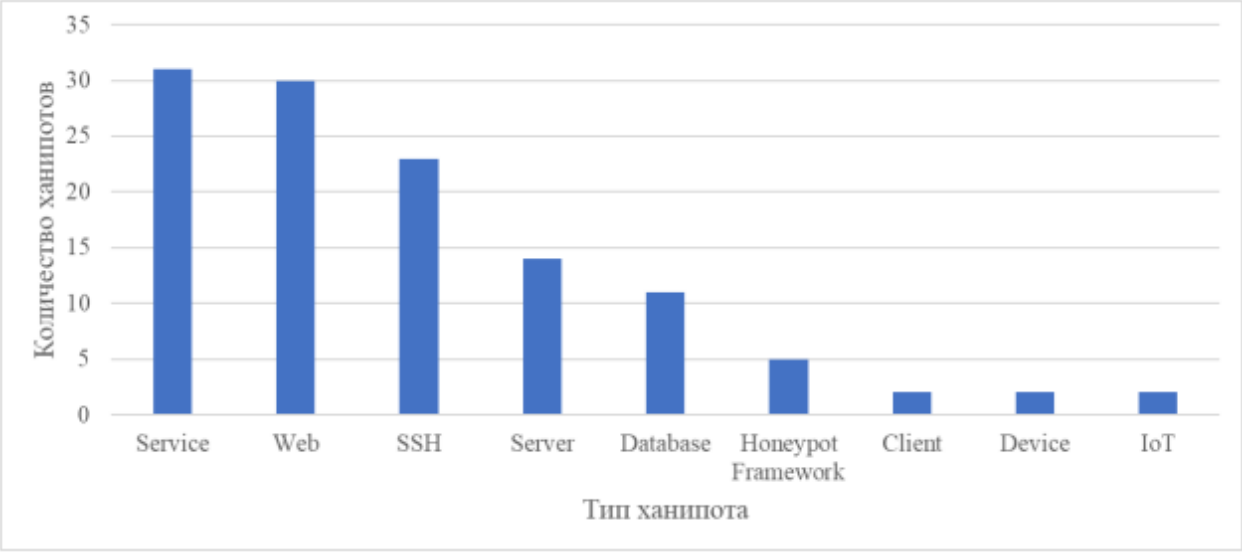


Рисунок 1 - Распределение honeypot по типам

При анализе статей был также рассмотрен потенциал использования honeypot-систем для обнаружения и изучения поведения веб-ботов. Работа «Защита веб-контента от несанкционированного автоматизированного сбора информации» [2] содержит исследование,

направленное на обзор методов защиты информации на веб-сайтах. Авторы Р.Ю. Демина, Д.Э. Шукралиева и Р.Р. Ференс объясняют разницу между легитимными и нелегитимными ботами и приводят статистику количества нежелательных ботов в сети.

Легитимные боты используются для сбора легальной информации согласно всем правилам сайтов и законов Российской Федерации, в то время как нелегитимные боты собирают информацию, которая содержит персональные данные пользователей, интеллектуальную собственность, финансово-значимый контент или информацию, предназначенную для нанесения ущерба веб-ресурсам. По данным этого исследования доля нежелательного роботизированного трафика в 2022 году составила 24% всего веб-трафика, что показывает значительный рост таких программ по сравнению с 2017 годом.

Авторы выделяют следующие методы защиты:

- CAPTCHA;
- динамическая смена идентификаторов значимых элементов;
- honeypot-системы;
- размещение информации картинками.

В работе «Применение технологии Honeypot для изучения поведения веб-роботов» [3] авторы А.А. Менщиков, Ю.А. Гатчин, А.В. Комарова более подробно объясняют значимость использования honeypot-систем для предотвращения угроз парсинга. Авторы вводят определение honeypot как ресурс-приманка, позволяющий не блокировать действия злоумышленников, и изучать стратегии их поведения. Конкретная реализация может кардинально отличаться в зависимости от целей и задач и представлять из себя как отдельный ресурс, так и динамическую систему, имитирующую реальный веб-сайт. Авторы работы провели исследование для выявления ботов и определения их поведения. Они пришли к выводу, что у большинства ботов отсутствует динамический анализ собранной информации. Также это исследование показывает, что honeypot-системы эффективны в обнаружении ботов и позволяют анализировать их поведение с целью дальнейшего противодействия им.

В анализе также рассмотрены работы, посвященные описанию технологии Honeypot. Работа «Использование Honeypot для привлечения внимания и поиска злоумышленников» [4] за авторством А.А. Сухова и В.В. Гармонова описывает все типы honeypot-систем, их достоинства и правила размещения. Авторы вводят еще одно определение honeypot. Honeypot – это системы безопасности, которые привлекают кибератаки и отслеживают их действия в защищенной, изолированной и контролируемой среде. Honeypot могут отвлекать потенциальные атаки от критически важных ресурсов и действовать как платформа для сбора информации об атакующих и их тактиках, техниках и процедурах. В исследовании авторы приводят типы honeypot и их определения:

1) Honeypot с низким уровнем взаимодействия. Он предназначен для имитации ограниченного числа сервисов, не сложен в реализации.

2) Honeypot с высоким уровнем взаимодействия. Он предлагает сложную, реалистичную среду, поскольку в состав ловушки могут входить дополнительные системы, базы данных и процессы.

3) Гибридный Honeypot. Данный вариант имитирует элементы прикладного уровня, но не имеет операционной системы.

В работе выделяю следующие достоинства honeypot-систем:

- раннее обнаружение атак;

- сбор разведывательной информации и анализ методов атак;
- обнаружение новых угроз;
- улучшение уровня безопасности и снижение рисков.

Авторы отмечают, что в мировой практике Honeypot используются для соблюдения специфических индустриальных требований. Их используют как часть комплексной и более широкой стратегии безопасности для обнаружения и реагирования на угрозы. Для того чтобы Honeypot был безопасен и эффективен в работе определяют следующие правила:

1) Он должен быть размещен в пределах сети, чтобы предложить потенциальным атакующим цель, при этом выглядя как обычная служба.

2) Он должен быть изолирован от остальной сети и не содержать важной информации.

3) Его журналы должны регулярно проверяться.

4) Honeypot должен управляться опытным персоналом и иметь регулярное обслуживание.

5) Он должен быть интегрируемым с другими мерами безопасности.

6) Юридические последствия использование Honeypot должны быть учтены, так как некоторые страны имеют строгие законы относительно мониторинга и перехвата связи.

Кроме отечественных научных работ были рассмотрены и зарубежные работы. Они показали, что интерес к технологии honeypot остается стабильно высоким. Это подтверждают значительное число публикаций, появившихся по данной теме за последние годы. В работе авторов A. Javadpour, F. Ja’fari, T. Taleb, M. Shojafar и C. Benzaïd «A comprehensive survey on cyber deception techniques to improve honeypot performance» [5] представлен масштабный обзор исследований в области honeypot, охватывающий последние два десятилетия. Целью работы стало систематизировать накопленные знания и обратить внимание на перспективные подходы и направления, которые пока остаются недостаточно изученными. Авторы детально описали существующие классификации honeypot-систем, а также различные методы обмана злоумышленников. К таким методам относятся расширенная имитация, манипуляции с базами данных и перенаправление трафика. Помимо этого, значительное внимание было уделено вопросам оптимизации приманок, моделированию атак и способам оценки эффективности honeypot-систем.

Данное исследование представляет ценность прежде всего как теоретическая база, на которую можно опираться при дальнейшей работе в этой области. Авторы предлагают ряд подходов к совершенствованию существующих систем защиты, что особенно актуально с точки зрения повышения точности и оперативности обнаружения вредоносной активности. Описанные методики могут послужить основой для создания более совершенных решений, способных выявлять угрозы на ранних стадиях и обеспечивать более надежную защиту.

В работе «A survey of contemporary open-source honeypots, frameworks, and tools» [6] авторы N. Ilg, P. Duplys, D. Sisejkovic и M. Menth обращаются к проблеме современных киберугроз и роли honeypot-технологий в противодействии им. На основе анализа большого числа исследований и решений с открытым исходным кодом авторы предлагают следующую классификацию honeypot-систем:

- как средство раннего обнаружения атак;
- как инструмент для сбора образцов вредоносного ПО;
- как способ изучения методов атак и уязвимостей.

Отдельно авторы подчеркивают роль технологий контейнеризации, которые существенно упрощают развертывание и автоматизацию honeypot-систем. Вместе с тем они указывают на сопутствующие риски, в частности на необходимость обеспечения безопасности самих контейнерных сред.

В рамках эксперимента, направленного на сравнительный анализ honeypot-решений, отдельное внимание было уделено подбору критериев оценки. Они выбирались исходя из того, насколько полно каждый из них способен отразить функциональную и техническую сторону системы в ситуации, когда существует реальная угроза безопасности. Для проведения анализа были определены следующие критерии:

1) Скорость обнаружения вторжений.
2) Уровень ложных срабатываний.
3) Объем и качество собираемых данных.
4) Устойчивость honeypot-системы.
5) Простота внедрения.
6) Производительность.

Данные критерии в совокупности охватывают наиболее значимые характеристики honeypot-решений, на которые следует обращать внимание при сравнительной оценке.

В ходе эксперимента использовалось несколько вариантов веб-форм, отличающихся механизмами защиты. Одна форма содержала скрытые honeypot-поля, предназначенные для обнаружения автоматического заполнения. Другие формы включали механизмы CAPTCHA, реализованные с использованием сервисов Google и Cloudflare. Схема тестового стенда изображена на Рисунке 2. Такое разделение позволяет сравнить эффективность различных подходов к защите форм от автоматизированных запросов в одинаковых условиях тестирования.

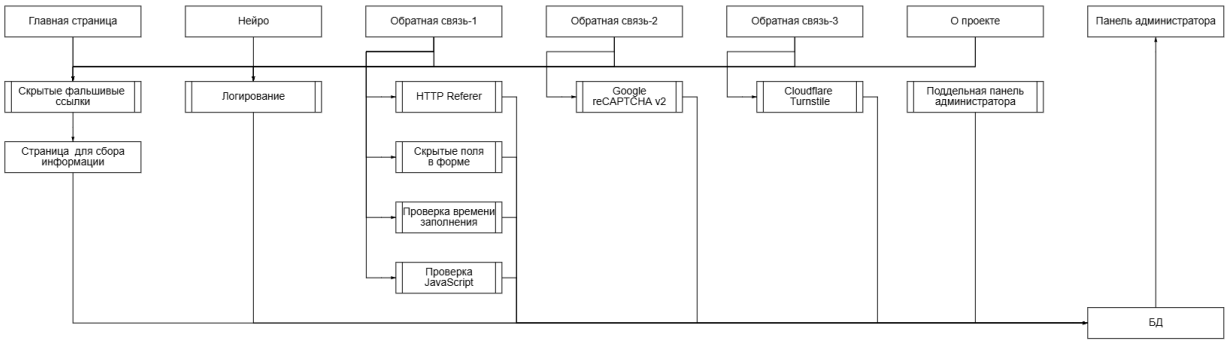


Рисунок 2 - Схема тестового стенда

Для проведения эксперимента были разработаны боты, имитирующие различные типы поведения. В тестовый набор вошли простые скриптовые боты, выполняющие базовые HTTP-запросы без эмуляции браузера, схемы которых изображены на Рисунке 3.

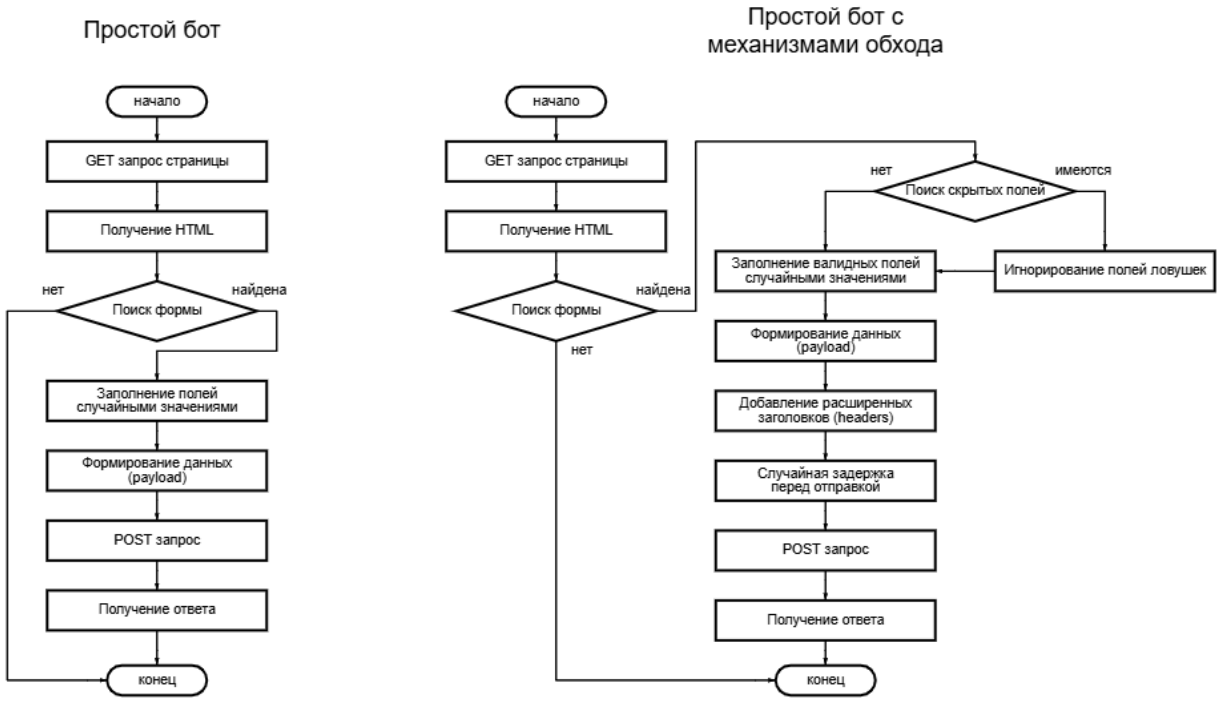


Рисунок 3 - Схемы простых ботов

И более сложные headless-боты, способные имитировать действия реального пользователя в браузерной среде, схемы которых изображены на Рисунке 4. Использование различных моделей поведения позволяет оценить устойчивость механизмов защиты веб-форм к автоматизированным атакам разного уровня сложности.

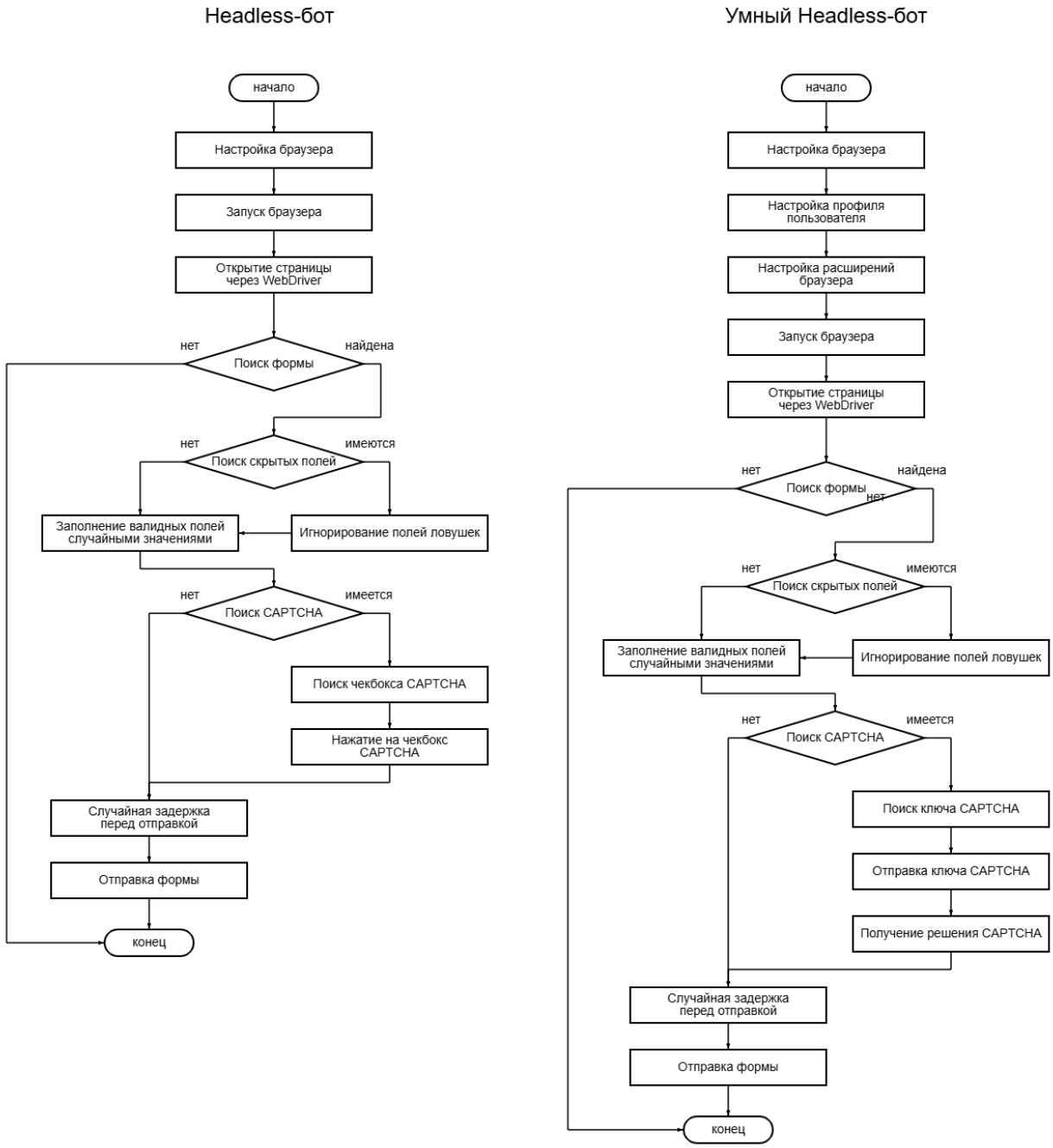


Рисунок 4 - Схемы headless-ботов

Для проведения тестирования был написан небольшой скрипт на языке программирования Python, который в автоматическом режиме запускал ботов для тестирования разных форм. В рамках эксперимента каждый бот запускался одну тысячу раз на каждую из разработанных форм. Такой подход позволил получить достаточный объем данных для последующего анализа эффективности honeypot-решений. Повторение большого количества попыток отправки форм также помогло снизить влияние случайных факторов и повысить достоверность результатов эксперимента. Таблица 1 содержит сравнительные результаты по 1000 запросам для пяти категорий ботов.

Таблица 1 - Сравнение частоты срабатывания ловушек на 1000 запросов

Тип ловушки	Простой бот	Простой бот с механизмами обхода	Headless-бот	Headless-бот с реальным профилем пользователя	Умный Headless-бот
HTTP Referer	1000	0	0	0	0
Скрытое поле (input)	1000	0	0	0	0
Скрытое поле (textarea)	1000	0	0	0	0
Проверка на выполнение JavaScript	1000	1000	0	0	0
Проверка времени заполнения формы	1000	98	21	19	0
Google reCAPTCHA v2	1000	1000	986	921	263
Cloudflare Turnstile	1000	1000	998	987	84

По результатам проведенного анализа можно утверждать, что простые honeypot-решения, показывают высокую эффективность только против самых простых ботов, а при переходе к более продвинутым ботам их эффективность падает до нуля. Из ловушек можно выделить проверку выполнения JavaScript, которая полностью блокирует простых ботов, но оказывается неэффективной против ботов, способных выполнять клиентские скрипты. А также выделить проверку времени заполнения формы. Видно, что её эффективность постепенно снижается по мере усложнения ботов. Таким образом, можно сделать вывод, что классические honeypot-решения демонстрируют слабую эффективность против современных ботов. Поэтому целесообразно рассмотреть более продвинутые механизмы защиты, такие как CAPTCHA.

Системы CAPTCHA продемонстрировали значительно более высокие результаты по сравнению с honeypot-решениями. В эксперименте были рассмотрены решения Google reCAPTCHA v2 и Cloudflare Turnstile. Оба механизма показали высокую эффективность против простых HTTP-ботов и ботов с базовыми механизмами обхода, обеспечивая 100% обнаружение автоматизированных запросов. Для обычного headless-бота Google reCAPTCHA v2 показал 98,6% блокировок, тогда как Cloudflare Turnstile показал 99,8%. Как показано на Рисунке 5, CAPTCHA пропустила лишь первые несколько запросов, после чего начала стабильно блокировать последующие попытки отправки формы.

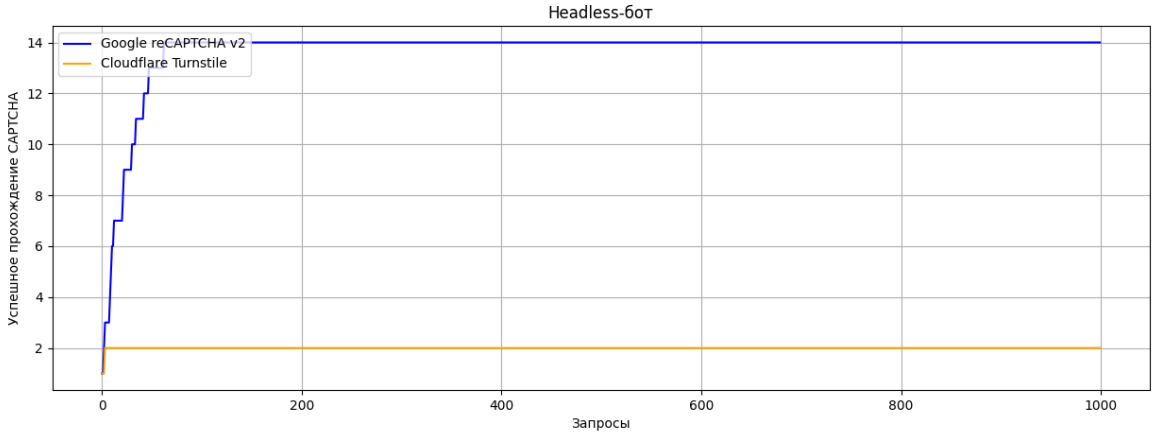


Рисунок 5 - Статистика прохождения CAPTCHA Headless-ботом

При использовании headless-бота с реальным браузерным профилем эффективность механизмов CAPTCHA несколько снижается. В данном случае Google reCAPTCHA v2 продемонстрировала уровень обнаружения 92,1%, тогда как Cloudflare Turnstile 98,7%. Это может быть связано с тем, что использование реального профиля браузера позволяет боту применять сохраненные cookies, параметры окружения и другие характеристики, приближающие его поведение к поведению реального пользователя. Как показано на Рисунке 6, данный тип бота успешно проходит проверку CAPTCHA на протяжении большего числа запросов, после чего эффективность прохождения постепенно снижается.

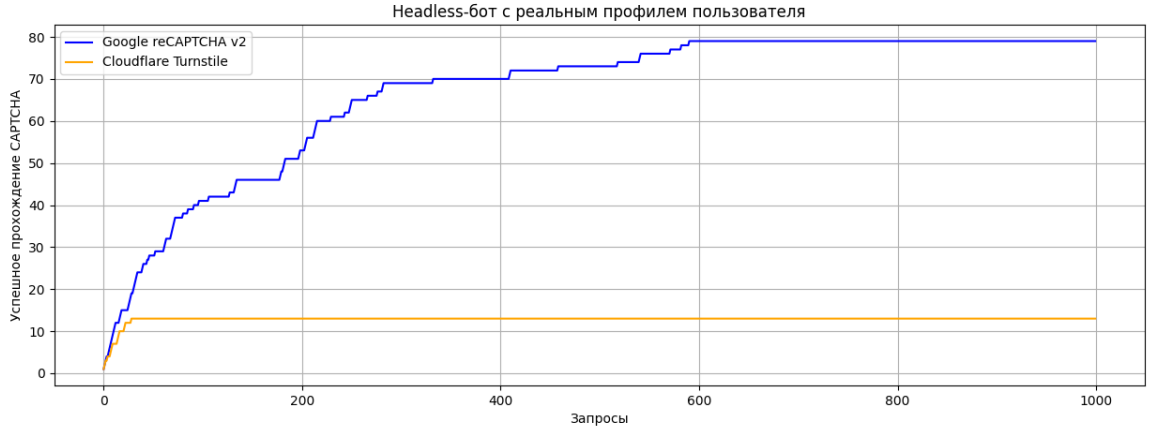


Рисунок 6 - Статистика прохождения CAPTCHA Headless-ботом с реальным профилем пользователя

Наиболее заметное снижение эффективности наблюдается при использовании умного headless-бота. В данном сценарии Google reCAPTCHA v2 зафиксировала 26,3% срабатываний, тогда как Cloudflare Turnstile зафиксировал только 8,4%. Это свидетельствует о том, что при реализации более сложных алгоритмов автоматизации, включающих имитацию пользовательского поведения, использование реального браузерного профиля, а также механизмы обхода CAPTCHA, современные защитные механизмы можно частично или полностью обойти. На Рисунке 7 представлены случаи успешного прохождения CAPTCHA данным типом бота.

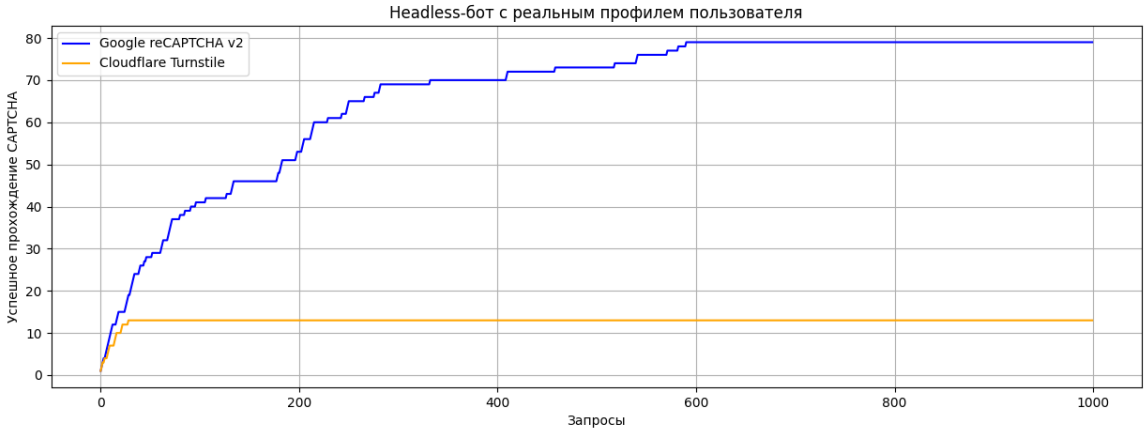


Рисунок 7 - Статистика прохождения CAPTCHA умным headless-ботом

Для оценки эффективности различных защитных механизмов веб-форм представлена Таблица 2, в которой обобщены результаты проведенного анализа.

Таблица 2 - Оценка эффективности ловушек для веб-формы

Тип ловушки	Скорость обнаружения	Ложные срабатывания	Объем и качество данных	Устойчивость	Простота внедрения	Производительность
HTTP Referer	Высокая	Низкие	Низкий	Низкая	Простое	Высокая
Скрытое поле (input)	Высокая	Низкие	Низкий	Низкая	Простое	Высокая
Скрытое поле (textarea)	Высокая	Низкие	Низкий	Низкая	Простое	Высокая
Проверка на выполнение JavaScript	Высокая	Низкие	Средний	Средняя	Простое	Высокая
Проверка времени заполнения формы	Средняя	Средние	Средний	Низкая	Простое	Высокая
Google reCAPTCHA v2	Средняя	Средние	Высокий	Высокая	Простое	Средняя
Cloudflare Turnstile	Средняя	Низкие	Высокий	Высокая	Простое	Средняя

Результаты показывают, что системы CAPTCHA обеспечивают значительно более высокий уровень обнаружения автоматизированных запросов по сравнению с простыми

honeypot-механизмами. При этом следует отметить, что поведенческие CAPTCHA, такие как Cloudflare Turnstile, способны эффективно перекрывать задачи, решаемые honeypot-ловушками, при этом не создавая дополнительных неудобств для пользователей, как это наблюдается при использовании Google reCAPTCHA v2. Это делает подобные решения наиболее предпочтительными для защиты веб-форм.

Список литературы

1. Сергановский А. А., Романович В. Н. Анализ ханипотов с открытым исходным кодом // Успехи в науке и образовании 2023 : сб. ст. II Междунар. науч.-исслед. конкурса. Пенза, 10 дек. 2023 г. Пенза : Наука и Просвещение, 2023. С. 76–80. EDN QPPZFM.
2. Демина Р.Ю., Шукралиева Д.Э., Ференс Р.Р. Защита веб-контента от несанкционированного автоматизированного сбора информации // Проблемы проектирования, применения и безопасности информационных систем в условиях цифровой экономики : материалы XXII Междунар. науч.-практ. конф. Ростов-на-Дону, 21–22 нояб. 2022 г. Ростов н/Д: Ростовский государственный экономический университет «РИНХ», 2022. С. 70–75. EDN AJLODZ.
3. Менщиков А.А., Гатчин Ю.А., Комарова А.В. Применение технологии honeypot для изучения поведения веб-роботов // Проблемы информационной безопасности : материалы VII Всерос. заоч. интернет-конф. Ростов-на-Дону, 20–21 февр. 2018 г. Ростов н/Д : АзовПринт, 2018. С. 45–48. EDN POPDGV.
4. Сухов А.А., Гармонов В.В. Использование Honeypot для привлечения внимания и поиска злоумышленников // Высокотехнологичное право: точка бифуркации: материалы V Междунар. науч.-практ. конф. : в 3 ч. Москва – Красноярск, 15–16 февр. 2024 г. Москва: Национальный исследовательский университет «МИЭТ», 2024. С. 32–37. EDN NGXWRZ.
5. Javadpour A., Ja'fari F., Taleb T., Shojafar M., Benzaid C. A comprehensive survey on cyber deception techniques to improve honeypot performance // Computers & Security. 2024. Vol. 140, no. 4. P. 103792. DOI: 10.1016/j.cose.2024.103792.
6. Ilg N., Duplys P., Sisejkovic D., Menth M. A survey of contemporary open-source honeypots, frameworks, and tools // Journal of Network and Computer Applications. 2023. Vol. 220. P. 103737. DOI: 10.1016/j.jnca.2023.103737. EDN KMKMPA.

References

1. Serganovsky A. A., Romanovich V. N. Analysis of open-source honeypots // Advances in Science and Education 2023: Coll. Art. of the II Int. Research Competition. Penza, December 10, 2023. Penza: Science and Education, 2023. pp. 76–80. EDN QPPZFM.
2. Demina R. Yu., Shukralieva D. E., Ferens R. R. Protecting web content from unauthorized automated collection of information // Problems of design, application, and security of information systems in the digital economy: Proc. of the XXII Int. Research and Practical Conf. Rostov-on-Don, November 21–22. 2022 Rostov n / D: Rostov State University of Economics "RINH", 2022. pp. 70-75. EDN AJLODZ.
3. Menshchikov A. A., Gatchin Yu. A., Komarova A. V. Using honeypot technology to study the behavior of web robots // Problems of information security: Proc. 7th All-Russian

correspondence internet conf. Rostov-on-Don, February 20-21, 2018 Rostov n / D: AzovPrint, 2018. pp. 45-48. EDN POPDGV.

4. Sukhov A. A., Garmonov V. V. Using honeypot to attract attention and find intruders // High-tech law: bifurcation point: Proc. V Int. scientific-practical. conf. : at 3 o'clock. Moscow - Krasnoyarsk, February 15-16, 2024. Moscow: National Research University of Electronic Technology, 2024. pp. 32-37. EDN NGXWRZ.
5. Javadpour A., Ja'fari F., Taleb T., Shojafar M., Benzaid C. A comprehensive survey on cyber deception techniques to improve honeypot performance // Computers & Security. 2024. Vol. 140, no. 4. p. 103792. DOI: 10.1016/j.cose.2024.103792.
6. Ilg N., Duplys P., Sisejkovic D., Menth M. A survey of contemporary open-source honeypots, frameworks, and tools // Journal of Network and Computer Applications. 2023. Vol. 220. P. 103737. DOI: 10.1016/j.jnca.2023.103737. EDN KMKMPA.