Электронный научный журнал "Математическое моделирование, компьютерный и натурный эксперимент в естественных науках" http://mathmod.esrae.ru/
URL статьи: **mathmod.esrae.ru/54-222**
Ссылка для цитирования этой статьи:
Перевертайло А.А., Хороводова Н.Ю., Кондратов Д.В. Разработка программного средства для деанонимизации пользователей на основе цифрового отпечатка браузера // Математическое моделирование, компьютерный и натурный эксперимент в естественных науках. 2025. №4

УДК 004.056.5 DOI:10.24412/2541-9269-2025-4-12-22

РАЗРАБОТКА ПРОГРАММНОГО СРЕДСТВА ДЛЯ ДЕАНОНИМИЗАЦИИ ПОЛЬЗОВАТЕЛЕЙ НА ОСНОВЕ ЦИФРОВОГО ОТПЕЧАТКА БРАУЗЕРА

Перевертайло А.А.[1], Хороводова Н.Ю.[2], Кондратов Д.В.[3]
[1] Саратовский государственный технический университет имени Гагарина Ю.А., Россия, Саратов, artemper2002@gmail.com
[2] Саратовский государственный технический университет имени Гагарина Ю.А., Россия, Саратов, khorovodovaniu@sstu.ru
[3] Саратовский государственный технический университет имени Гагарина Ю.А., Россия, Саратов, Институт проблем точной механики и управления Российской академии наук (ИПТМУ РАН), г. Саратов, Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского, kondratovdv@yandex.ru

DEVELOPMENT OF SOFTWARE FOR USER DE-ANONYMIZATION BASED ON DIGITAL BROWSER FINGERPRINTS

Perevertaylo A.A.[1], Khorovodova N.Yu.[2], Kondratov D.V.[3]
[1]Yuri Gagarin State Technical University of Saratov, Russia, Saratov, artemper2002@gmail.com
[2]Yuri Gagarin State Technical University of Saratov, Russia, Saratov, khorovodovaniu@sstu.ru
[3]Yuri Gagarin State Technical University of Saratov, Russia, Saratov, Institute of Precision Mechanics and Control of the Russian Academy of Sciences, Saratov, Russia; Saratov State University, Saratov, Russia kondratovdv@yandex.ru

Аннотация. В статье исследуются современные методы деанонимизации интернет-пользователей, применяющих VPN и другие средства сокрытия данных. Рассмотрена технология browser fingerprinting, основанная на сборе уникальных характеристик браузера и устройства, позволяющая идентифицировать пользователей с высокой точностью. Приведены результаты тестирования разработанного модуля, подтверждающие его устойчивость к изменениям конфигурации и эффективность по сравнению с cookie-ориентированными системами.

Ключевые слова: информационная система, безопасность информационной системы, отпечаток браузера, анонимность, деанонимизация пользователя, аналитика данных

Abstract. The article examines modern methods of internet user de-anonymization when VPN and other anonymization tools are applied. It focuses on browser fingerprinting technology, which collects unique browser and device characteristics to ensure high-accuracy identification. The paper presents testing results of the developed module, demonstrating its resilience to configuration changes and higher efficiency compared to cookie-based systems.

Keywords: information system, information system security, browser fingerprint, anonymity, user deanonymization, data analytics

Современное развитие цифровых технологий сопровождается не только ростом комфорта и доступности онлайн-сервисов, но и повышением интереса к вопросам конфиденциальности, анонимности и информационной безопасности. Всё больше пользователей стремятся скрыть своё присутствие в сети, применяя средства анонимизации – прежде всего VPN-сервисы, которые позволяют изменить IP-адрес и затруднить отслеживание активности. Эти технологии находят широкое применение как среди обычных пользователей, стремящихся защитить личную информацию, так и среди злоумышленников, использующих анонимность в преступных целях.

Деанонимизация в интернете: методы, задачи и перспективы

В условиях, когда стандартные механизмы отслеживания пользователей становятся всё менее эффективными, возрастает потребность в разработке новых инструментов деанонимизации.[1-7]

Деанонимизация в интернете охватывает методы, средства и технологии, позволяющие идентифицировать пользователей, скрывающих свою личность с помощью различных инструментов анонимизации.

Одним из ключевых методов деанонимизации является технология "отпечатка браузера" (browser fingerprinting), которая базируется на сборе уникальных характеристик браузера и устройства пользователя.

Предметная область деанонимизации имеет широкую сферу применения. В корпоративной и государственной безопасности деанонимизация помогает предотвращать мошенничество, защищать данные и отслеживать подозрительную активность в сети. Компании используют эти технологии для защиты корпоративной информации, предотвращения кражи данных, а также в целях соблюдения нормативных требований. Государственные структуры и правоохранительные органы применяют методы деанонимизации для расследования киберпреступлений, противодействия финансированию терроризма и борьбы с дезинформацией.

Технология браузерного отпечатка основана на использовании уникального набора характеристик, присущего конкретной комбинации браузера, операционной системы и аппаратной конфигурации устройства. В отличие от IP-адреса, который можно быстро подменить с помощью VPN,

отпечаток браузера позволяет с высокой точностью идентифицировать пользователя даже при попытках анонимизации.

В настоящей работе рассматриваются результаты создания программного средства, способного деанонимизировать VPN-подключения за счёт сопоставления информации об отпечатках браузера и IP-адресах, собираемой с различных веб-ресурсов. Система представляет собой модуль, который может быть внедрён в любой сайт и функционирует как централизованный сервис анализа цифровых следов.

Исходя из вышеизложенного в созданном программном модуле реализованы следующие механизмы.

Механизм сбора данных:

- Определение IP-адреса пользователя.
- Сбор информации об аппаратных и программных характеристиках браузера (User-Agent, WebGL, Canvas, шрифты, плагины и т. д.).
- Обнаружение использования прокси и VPN-сервисов.

Механизм центрального хранилища данных:

- Формирование базы данных для хранения отпечатков браузеров.
- Разработка алгоритмов сопоставления данных с разных сайтов.
- Обеспечение защиты собранной информации.

Механизм анализа и выявления VPN-подключений:

- Определение закономерностей в изменениях IP-адресов и отпечатков браузеров.
- Составление отчётов о пользователях с возможными VPN-подключениями.

Механизм API для интеграции с сайтами:

- Предоставление удобного способа подключения модуля к различным веб-ресурсам.
- Разработка методов передачи данных в централизованную систему.

Тестирование и оптимизация:

- Проверка точности определения VPN-подключений.
- Оптимизация работы модуля для минимального влияния на производительность веб-ресурсов.

Преимущества над аналогичными сервисами

Предлагаемое решение является аналитическим инструментом, оно схоже по функциональности с Google analytics и Яндекс метрика, но для сбора информации о пользователях эти инструменты используют cookie файлы, что не позволяет корректно определять необходимого пользователя для его деанонимизации, так как у этого способа есть ряд проблем:

- Начиная с конца 2019 года браузеры стали активно ограничивать использование файлов cookie — как сторонних (third-party), так и собственных (first-party). Первые создаются внешними сервисами, код которых владелец сайта добавляет к себе (например, чаты, всплывающие окна, кнопки соцсетей). Вторые формируются самим сайтом (доменом), на котором находится пользователь, и в основном применяются для аналитики.[8]
- Срок действия cookie часто сокращается до менее чем 10 дней из-за встроенных ограничений браузеров, механизмов интеллектуального отслеживания, обновлений iOS и других факторов.
- Пользователь может самостоятельно удалить историю посещений вместе с cookie.
- Дополнительно он может использовать блокировщики рекламы и расширения, препятствующие сбору данных аналитическими счетчиками.
- Антивирусные программы, установленные на компьютере, также способны мешать корректному функционированию браузера и cookie.
- При использовании режима «Инкогнито» файлы cookie, история просмотров, данные сайтов и введённая в формы информация не сохраняются.

Всё это ведёт к ряду последствий:

- Несоответствие данных: например, Google Analytics 4 может быть заблокирован браузером, тогда как Яндекс.Метрика останется доступной. В итоге пользователь отразится только в одной системе.
- Сокращение цепочек взаимодействия: при повторном визите пользователя создаётся новый cookie и новая карточка в аналитике, поэтому связать прошлые действия с текущими уже нельзя.
- Эффективность рекламных кампаний нередко оценивается по модели последнего клика (Last Click) или по непрозрачным алгоритмам на основе данных («чёрный ящик»).
- Маркетологи смещают акцент на сбор контактных данных (телефоны, e-mail) для ретаргетинга и работы с клиентской базой.
- Аналитика всё больше основывается на идентификации пользователей через User ID (авторизация), а не через Client ID.

Если пользователь вообще отказывается от отслеживания, запрещает сбор файлов cookie, то поведенческие данные о нем станут недоступны, и тогда в дело вступает так называемое Моделирование. В этом случае система использует сведения о похожих людях, которые дали тому же ресурсу согласие на использование файлов cookie, чтобы смоделировать недостающую информацию, что не позволит нам корректно деанонимизировать пользователей.

Основываясь на этих данных, для разработки была выбрана библиотека FingerprintJS, чтобы получать данные о пользователях.

Принцип работы и особенности выбранной библиотеки

FingerprintJS - это быстрая библиотека отпечатков браузера. По умолчанию используется Murmur hashing и возвращает 32-битное целое число. Стандартную функцию хеширования можно легко заменить, что делает ее гибкой в использовании.[9]

Браузеру запрашивается его строка агента, разрешение экрана и глубина цвета, установленные плагины с поддерживаемыми типами mime, смещение часового пояса и другие возможности, такие как локальное хранилище и хранилище сеансов, и другие схожие параметры. Затем эти значения передаются через функцию хеширования для создания отпечатка.

FingerprintJS, вообще, не использует cookie. Никакая информация не сохраняется на жестком диске компьютера, где установлен браузер. Работает в инкогнито режиме, потому что в принципе не использует хранение на жестком диске.[9]

Отдельным шагом идёт проверка поддержки HTML5-технологий. У каждого браузера этот набор различается, поэтому FingerprintJS последовательно опрашивает все доступные возможности: какие функции доступны, какие отсутствуют, и в какой степени они реализованы. Эти результаты тоже становятся частью цифрового отпечатка.[9]

Также используется Canvas API для проверки отрисовки шрифтов. Поскольку результат зависит от процессора, видеокарты, драйверов и системных библиотек, одинаковые изображения в разных браузерах преобразуются в разные байтовые массивы. Эти данные добавляются к итоговому отпечатку.

Возможные проблемы при сборе отпечатков

1. UserAgent. У современных браузеров UserAgent меняется очень часто, А это означает, что UserAgent может влиять на итоговый отпечаток.

2. IPhone, IPad и другие продукты Apple. Дело в том, что они очень униформенные, одинаковые с точки зрения аппаратной реализации. У них у всех одинаковые процессоры, а это значит, что точность идентификации для продуктов Apple будет ниже.

Для повышения точности используется fuzzy hashing (нечёткое хэширование), при котором небольшие изменения входных данных не влияют на итоговый отпечаток.

В FingerprintJS 2 учитываются установленные шрифты: разные программы (например, Office или Adobe PDF) добавляют свои наборы, что позволяет различать даже схожие компьютеры.

Ещё один метод — WebGL Fingerprint, развитие Canvas Fingerprint. Он строит 3D-объекты с эффектами и преобразует их в байтовый массив, зависящий от видеокарты, драйверов и ОС. Эти различия делают отпечаток

уникальным и позволяют точнее идентифицировать устройства, включая iOS (поддержка WebGL появилась с версии 8.1).[9]

.Все эти данные в совокупности позволяют получить точность определения порядка 94-95%.

Принцип работы сервиса

На рисунке 1 показана сквозная цепочка прохождения данных от конечных пользователей до центрального хранилища и административного интерфейса. Каждый пользователь открывает страницу одного из подключённых веб-сайтов . В момент обращения браузер передаёт сайту свой внешний IP-адрес и набор характеристик окружения, на основании которых формируется цифровой отпечаток. Клиентский скрипт сайта упаковывает эти сведения в Data и пересылает их на единый backend-сервер, где располагаются:

• слой приёма API,

• база данных журналов.

Центральный сервер собирает информацию сразу с нескольких сайтов, устраняя дубли и связывая события с ранее сохранёнными отпечатками. Авторизованный оператор, обладающий ролью Admin, подключается к тому же серверу и через web-интерфейс получает log info - журналы посещений, которые уже содержат сопоставления «пользователь ↔ IP ↔ отпечаток ↔ сайт». Таким образом схема отражает распределённую архитектуру: сбор первичных данных происходит на периферийных сайтах, а обработка, хранение и аналитика - в централизованном контуре под контролем администратора.

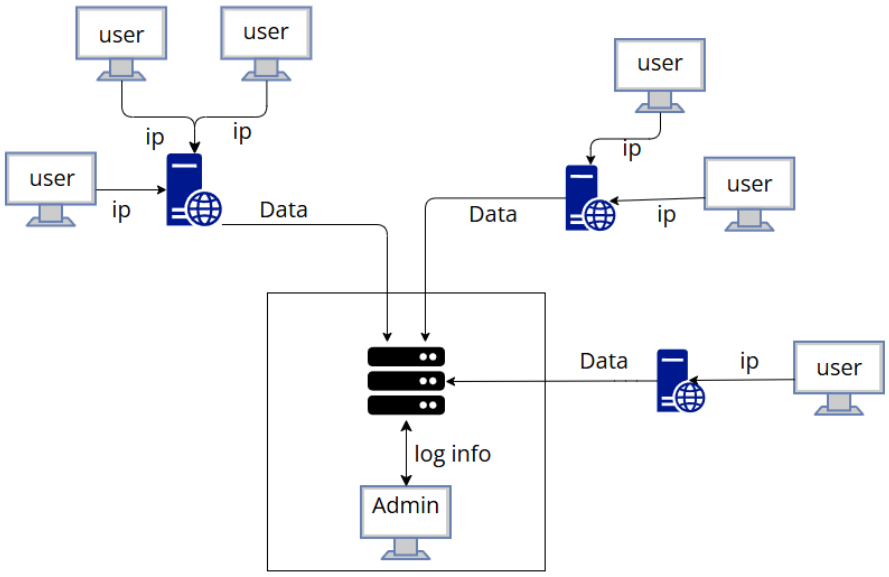


Рисунок 1. Структура работы сервера с внешними пользователями и сбор данных

Тестирование работоспособности сервиса

Для тестирования разработанного сервиса была разработана тестовая среда, представленная на рисунке 2.

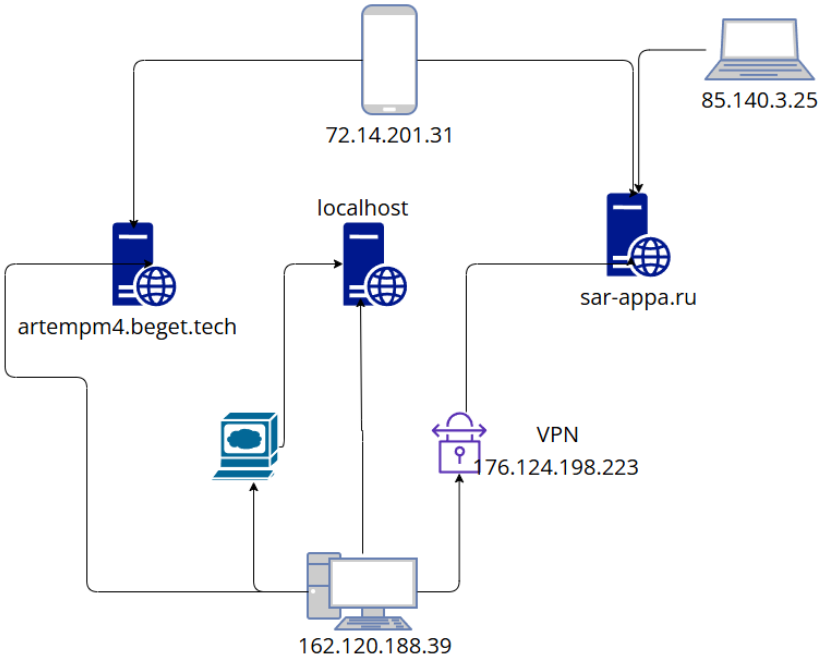
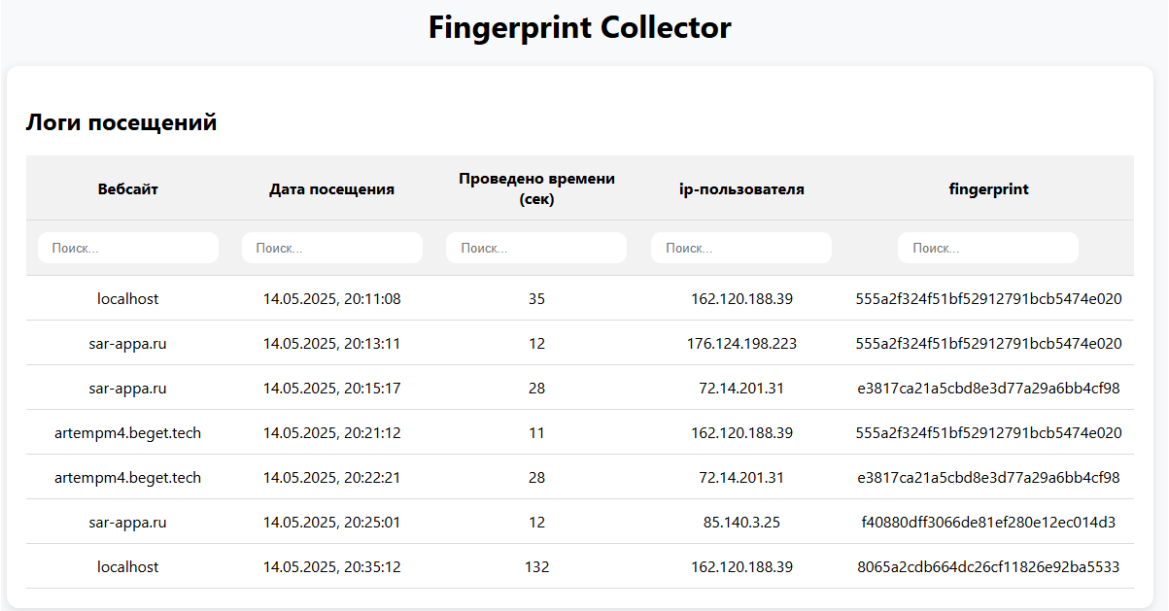


Рисунок 2. Тестовая среда

Испытания проводились на трёх площадках: локальный стенд (localhost), облачный хостинг artempm4.beget.tech и публичный домен sar-appa.ru. Клиентская часть тестировалась с ПК (162.120.188.39), ноутбука (85.140.3.25) и смартфона (72.14.201.31). Для моделирования анонимного трафика ПК дополнительно подключался через VPN (76.124.198.223), что позволило оценить устойчивость дактилоскопии и подтвердить различение устройств даже при смене конфигурации.

На рисунке 3 представлена выборка логов, полученная в результате контрольных испытаний. Для каждого визита зафиксированы домен сайта, метка времени, длительность сессии, реальный IP-адрес и хэш-отпечаток браузера (visitorId).



Fingerprint Collector

Логи посещений

Вебсайт	Дата посещения	Проведено времени (сек)	ip-пользователя	fingerprint
Поиск...	Поиск...	Поиск...	Поиск...	Поиск...
localhost	14.05.2025, 20:11:08	35	162.120.188.39	555a2f324f51bf52912791bcb5474e020
sar-appa.ru	14.05.2025, 20:13:11	12	176.124.198.223	555a2f324f51bf52912791bcb5474e020
sar-appa.ru	14.05.2025, 20:15:17	28	72.14.201.31	e3817ca21a5cbd8e3d77a29a6bb4cf98
artempm4.beget.tech	14.05.2025, 20:21:12	11	162.120.188.39	555a2f324f51bf52912791bcb5474e020
artempm4.beget.tech	14.05.2025, 20:22:21	28	72.14.201.31	e3817ca21a5cbd8e3d77a29a6bb4cf98
sar-appa.ru	14.05.2025, 20:25:01	12	85.140.3.25	f40880dff3066de81ef280e12ec014d3
localhost	14.05.2025, 20:35:12	132	162.120.188.39	8065a2cdb664dc26cf11826e92ba5533

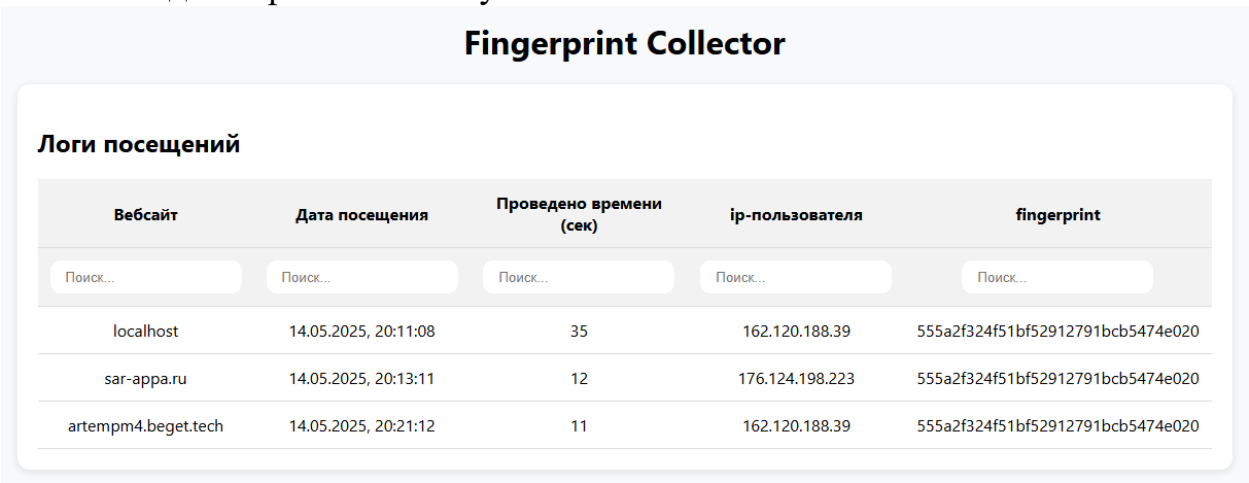
Рисунок 3. Результаты испытаний

Анализ результатов тестирования

Анализ этих строк позволяет сделать три ключевых наблюдения.

1. Стабильность отпечатка при смене IP.

Посещения, изображенные на рисунке 4 относятся к одному и тому же стационарному компьютеру (исходный адрес 162 .120 .188 .39), вторая строка — к тому же устройству, но через VPN-тоннель 76 .124 .198 .223. Во всех трёх случаях хэш остаётся неизменным, а значит, даже полная подмена внешнего IP не влияет на идентификацию: система уверенно связывает разные сетевые события с одним физическим субъектом.



Fingerprint Collector

Логи посещений

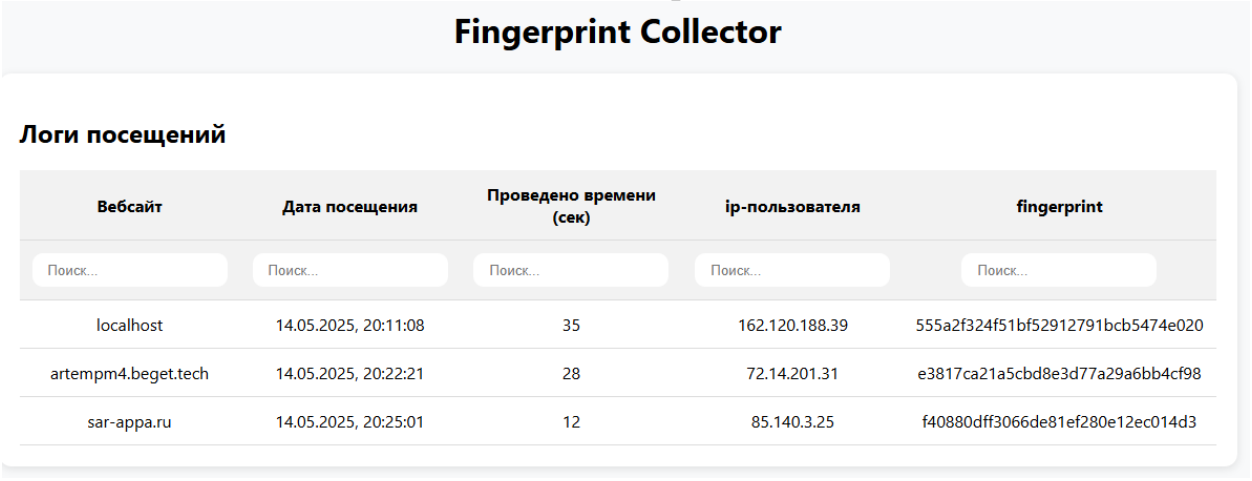
Вебсайт	Дата посещения	Проведено времени (сек)	ip-пользователя	fingerprint
Поиск...	Поиск...	Поиск...	Поиск...	Поиск...
localhost	14.05.2025, 20:11:08	35	162.120.188.39	555a2f324f51bf52912791bcb5474e020
sar-appa.ru	14.05.2025, 20:13:11	12	176.124.198.223	555a2f324f51bf52912791bcb5474e020
artempm4.beget.tech	14.05.2025, 20:21:12	11	162.120.188.39	555a2f324f51bf52912791bcb5474e020

Рисунок 4. Стабильность отпечатка при смене IP

2. Отслеживание пользователя при смене устройства.

Строки, изображенные на рисунке 5 демонстрируют обращения со стационарного ПК (IP 162.120.188.39), смартфона (IP 72.14.201.31), и ноутбука

(IP 85.140.3.25). У каждого устройства свой уникальный хэш, поэтому никакой ошибочной «склейки» с основным ПК не происходит.



Fingerprint Collector

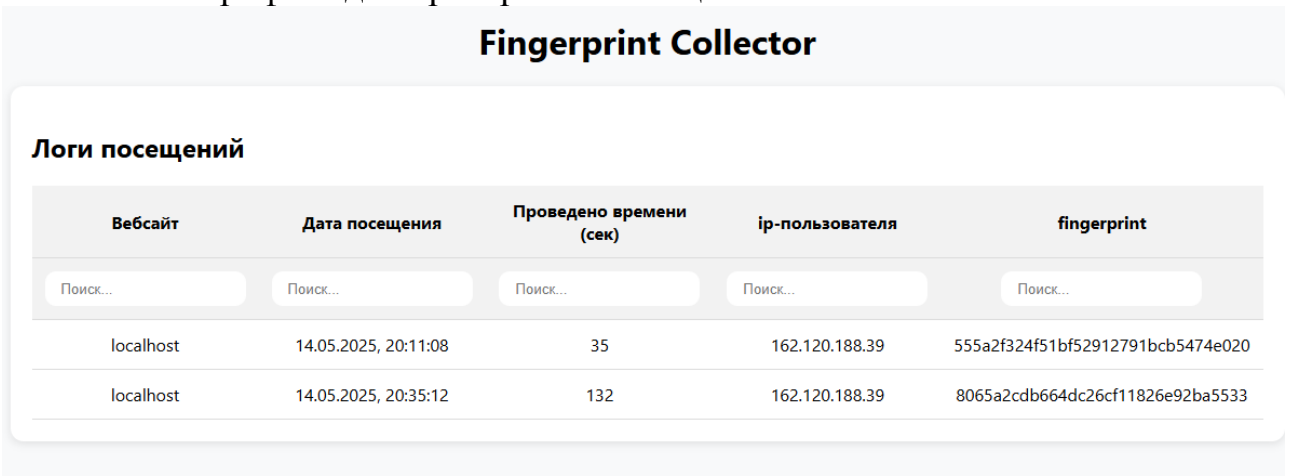
Логи посещений

Вебсайт	Дата посещения	Проведено времени (сек)	ip-пользователя	fingerprint
Поиск...	Поиск...	Поиск...	Поиск...	Поиск...
localhost	14.05.2025, 20:11:08	35	162.120.188.39	555a2f324f51bf52912791bcb5474e020
artempm4.beget.tech	14.05.2025, 20:22:21	28	72.14.201.31	e3817ca21a5cbd8e3d77a29a6bb4cf98
sar-appa.ru	14.05.2025, 20:25:01	12	85.140.3.25	f40880dff3066de81ef280e12ec014d3

Рисунок 5. Отслеживание пользователя при смене устройства

3. Чувствительность отпечатка к конфигурации браузера.

На рисунке 6 представлены строки посещений, между которыми изменялся набор расширений и разрешение монитора на том же компьютере. В результате первичный хэш сменился. Такой дрейф показывает, что текущий набор компонент FingerprintJS учитывает параметры, подверженные модификации. Если цель – свести количество ложных срабатываний к минимуму, следует расширить перечень собираемых характеристик, например, увеличить количество шрифтов для проверки с помощью Canvas API.



Fingerprint Collector

Логи посещений

Вебсайт	Дата посещения	Проведено времени (сек)	ip-пользователя	fingerprint
Поиск...	Поиск...	Поиск...	Поиск...	Поиск...
localhost	14.05.2025, 20:11:08	35	162.120.188.39	555a2f324f51bf52912791bcb5474e020
localhost	14.05.2025, 20:35:12	132	162.120.188.39	8065a2cdb664dc26cf11826e92ba5533

Рисунок 6. Чувствительность отпечатка к изменениям

Рекомендации по применению

Для успешного внедрения и дальнейшего развития разработанной системы следует:

Выполнять интеграцию клиентского скрипта на стороне веб-сайта через собственный поддомен или корпоративный CDN, что исключит зависимость от сторонних хостов и снизит вероятность блокировки расширениями браузера.

Вызов библиотеки FingerprintJS рекомендуется инициировать после события DOMContentLoaded, а передачу первичных данных осуществлять исключительно по HTTPS, чтобы минимизировать влияние на производительность страниц и сразу обеспечить шифрование канала.

В области информационной безопасности необходимо использовать mTLS для всех внутренних сервисных взаимодействий, хранить поля ip, visitorId и hashComponents в зашифрованном виде (AES-256, master-key в HSM) и вести неизменяемый аудит всех запросов привилегированных ролей; помимо этого, следует получить явное согласие пользователей на обработку отпечатков и реализовать механизмы «права на забвение», что позволит соответствовать требованиям GDPR и ФЗ-152.

Перспективы развития продукта связаны с расширением набора характеристик и применением нейросетевых моделей: внедрение «размытого» сравнения, основанного на обученной нейронной сети, способно динамически пересчитывать веса компонент отпечатка и тем самым повышать устойчивость к изменениям пользовательской конфигурации. Также это позволит вести риск-скоринг с учётом аномальных факторов (резкой смены геолокации, прокси, Canvas-сигнатур), а экспорт данных в SIEM-системы обеспечит сквозной анализ инцидентов безопасности на уровне корпоративной инфраструктуры.

Реализация перечисленных мер не только усилит защищённость решения, но и значительно повысит надёжность идентификации пользователей в реальных условиях эксплуатации.

Разработанный сервис будет полезен как владельцам веб-сайтов, заинтересованным в сборе статистики о действиях пользователей на своих ресурсах, так и специалистам в сфере кибербезопасности. Для первых он станет удобным инструментом аналитики, а его массовое внедрение на большом числе сайтов позволит службам кибербезопасности использовать сервис как дополнительное средство, ускоряющее расследование киберпреступлений.

Литература

1. Команда ishosting. Что такое Fingerprinting и в чем его роль в онлайн-пространстве / ishosting [Электронный ресурс]: 25 мар 2024 г. — URL: https://blog.ishosting.com/ru/what-is-fingerprinting (дата обращения: 05.04.2025).
2. Радайкин М. Ф. Кратко о проблеме анонимности в сети Интернет [Электронный ресурс]: CyberLeninka — URL: https://cyberleninka.ru/article/n/kratko-o-probleme-anonimnosti-v-seti-internet/viewer (дата обращения: 26.03.2025).
3. Проблемы и технологии защиты информации [Электронный ресурс]: Habr — URL: https://habr.com/ru/companies/oleg-bunin/articles/321294/ (дата обращения: 05.04.2025).

4. Шармаев В. И., Карпухин Е. О., Сидорин С. Ю., Жердев А. А. Деанонимизация пользователя веб-ресурса с применением технологии формирования отпечатка браузера // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. — 2022. — № 6. — С. 161–167.
5. Проблемы и технологии защиты информации [Электронный ресурс]: Habr — URL: https://habr.com/ru/companies/oleg-bunin/articles/321294/ (дата обращения: 05.04.2025)
6. Отказ от third-party cookies: что делать бизнесу и маркетологам [Электронный ресурс] — URL: https://garpun.com/otkaz-ot-third-party-cookies (дата обращения: 31.03.2025)
7. Колисниченко Д. Н. Анонимность и безопасность в Интернете. От «чайника» к пользователю. — СПб. : БХВ-Петербург, 2012. — 240 с.
8. Hacker Place. Почему TOR небезопасен? Сущность cookies и fingerprint. [Электронный ресурс]: 06 июн 2019 г. — URL: https://telegra.ph/Pochemu-TOR-nebezopasen-Sushchnost-cookies-i-fingerprint-06-06 (дата обращения: 08.04.2025).
9. Fingerprint.com [Электронный ресурс] — URL: https://fingerprint.com/ (дата обращения: 21.03.2025)