УДК 004.056.52

Н.С. Афанасьева, П.С. Ложников

Наборы данных, используемые для определения ботов на основании движения компьютерной мыши

Рассматриваются обнаружения ботов на основании динамических характеристик движения курсора компьютерной мыши, а также базовые динамические характеристики и полученные на их основании 150 добавочных. Приведены общедоступные наборы данных, которые используются для нахождения ботов по особенностям перемещения курсора компьютерной мыши. Представлены самостоятельно созданные датасеты BOT1, используемые для обучения системы. Для подтверждения применимости этих датасетов проведено снижение размерности и сравнение распределения ключевых динамических характеристик наборов.

Ключевые слова: бот, вредоносный бот, движение курсора компьютерной мыши, трек, наборы данных динамики мыши.

DOI: 10.21293/1818-0442-2024-27-3-118-124

В настоящее время боты (программы, автоматически выполняющие заранее настроенные повторяющиеся задачи с использованием интерфейсов, предназначенных для людей [1]) стали неотъемлемой частью повседневной жизни человека.

Продуктивные боты помогают улучшить индексацию в браузерах, обеспечить вовлеченность клиентов, масштабировать операции.

Вредоносные боты представляют собой автономные интеллектуальные продукты, разработанные с целью мошенничества, причинения вреда конечным пользователям или организациям в целом (парсинг веб-страниц, конкурентный анализ данных, скальпинг, скрейпинг личных и финансовых данных, попытки грубого входа в систему, атаки типа «отказ в обслуживании», мошенничество с цифровой рекламой, рассылка спама, мошенничество с транзакциями и другие подобные действия) [3].

Существует несколько схожих терминов, встречающихся в работах, посвященных вредоносным ботам, среди них – Bad bot [4, 5], плохие боты [6], злонамеренные боты [7], malicious bot [1, 8, 9], деструктивные боты [10]. В данном исследовании понятие вредоносного бота используется как синоним вышеперечисленных.

Вредоносные боты развили свои эффективные методы уклонения от традиционных решений безопасности, таких как брандмауэр веб-приложений (WAF) и CAPTCHA на основе имитации действий реальных пользователей.

Согласно последним отраслевым отчетам [10], трафик вредоносных ботов растет пятый год подряд. В 2023 г. их доля составила 32% от всего трафика, что на 1,8% больше, чем годом ранее. Доля продуктивных ботов также выросла – с 17,3 до 17,6%. В совокупности за 2023 г. доля интернет-трафика, сгенерированного ботами, составляет 49,6% от общего мирового трафика.

Борьба с ботами является комплексной проблемой, затрагивающей различные сферы и функции. Возможность ботов выполнять действия с гораздо большей скоростью и частотой, чем обычными пользователями, делает их эффективным инструментом для совершения вредоносных действий и атак.

Open Web Application Security Project (OWASP) – сообщество, включающее в себя ИТ-корпорации, образовательные организации, ИТ-специалистов, ориентированное на улучшение защищенности веб-приложений, выделило автоматизированные угрозы веб-приложениям с помощью ботов в отдельный класс, вне широко известного и применяемого проекта Топ-10 OWASP [11].

Веб-приложения подвергаются нежелательному автоматическому использованию каждый день. Эти события зачастую связаны с излишним использованием допустимой функциональности, а не с попыткой эксплуатации неустраненных уязвимостей.

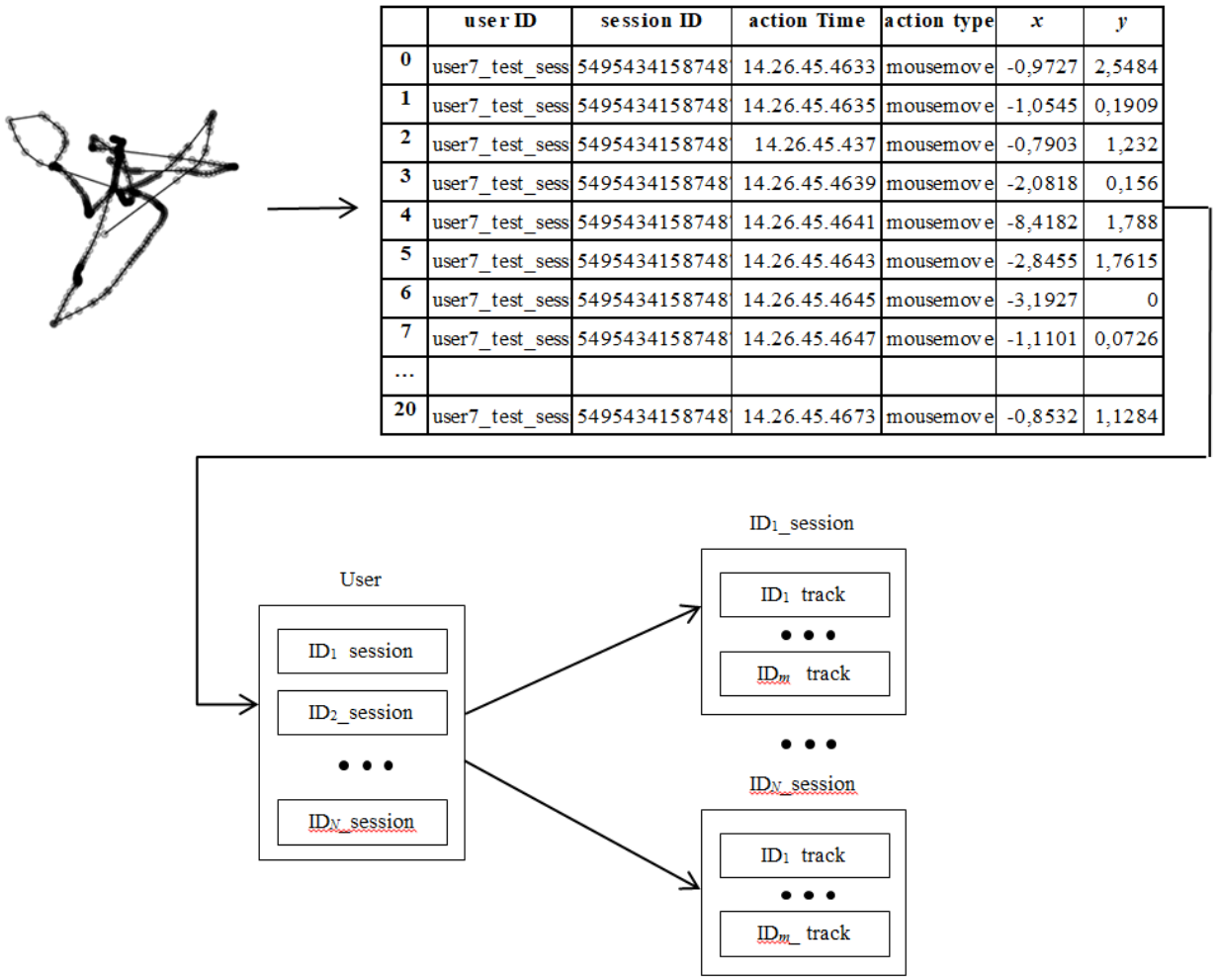
В качестве дополнительного инструмента для борьбы с автоматизированными угрозами, инициируемыми ботами, можно рассмотреть детектирование ботов на основании движения мыши. В отличие от динамики нажатия клавиш, анализ перемещения курсора мыши не относится к категории персональных данных пользователя, т.е. нет возможности зафиксировать, например, логин и пароль, вводимые с клавиатуры. Также в веб-средах использование мыши происходит чаще, чем клавиатуры [12].

Хотя некоторые исследования уже были проведены в вопросах детектирования ботов, не существует продукта, который позволял бы полностью решить эту проблему, поэтому разработка новых подходов к анализу динамики мышей остается областью, требующей дальнейших исследований. Важным ограничением для развития работ в области распознавания ботов по двигательной активности курсора компьютерной мыши является недоступность наборов данных. Общедоступные наборы данных движения мыши позволяют объективно сравнивать различные методы решения конкретных проблем.

Определение ботов на основе динамики мыши

В этой статье представлены характеристики, полученные на основании действий компьютерной мыши, а также наборы данных, используемые для обучения модели распознавания ботов.

Движение мыши пользователем представлено как набор сессий. Каждая сессия – это комплект треков. Трек представляет собой действия пользователя: перемещение мыши и нажатие клавиш. Причем каждый новый трек начинается, когда пользователь сделал паузу между своими действиями не менее чем на 3 с. Процесс анализа движения компьютерной мыши пользователем на треки представлен на рис. 1.



	user ID	session ID	action Time	action type	x	y
0	user7_test_sess	5495434158748	14.26.45.4633	mousemove	-0,9727	2,5484
1	user7_test_sess	5495434158748	14.26.45.4635	mousemove	-1,0545	0,1909
2	user7_test_sess	5495434158748	14.26.45.437	mousemove	-0,7903	1,232
3	user7_test_sess	5495434158748	14.26.45.4639	mousemove	-2,0818	0,156
4	user7_test_sess	5495434158748	14.26.45.4641	mousemove	-8,4182	1,788
5	user7_test_sess	5495434158748	14.26.45.4643	mousemove	-2,8455	1,7615
6	user7_test_sess	5495434158748	14.26.45.4645	mousemove	-3,1927	0
7	user7_test_sess	5495434158748	14.26.45.4647	mousemove	-1,1101	0,0726
…						
20	user7_test_sess	5495434158748	14.26.45.4673	mousemove	-0,8532	1,1284

Рис. 1. Процесс создания пользовательской модели

Характеристики мыши

В качестве динамических характеристик компьютерной мыши будем рассматривать *x*- и *y*-координаты точек трека, значения времени для каждого события мыши, набор базовых характеристик и набор добавочных характеристик.

К базовым характеристикам относятся длина пройденной траектории (*s*), угол между касательной траектории и осью *x* (θ_i) горизонтальная (v_x), вертикальная (v_y) и общая скорость (v), угловая скорость (ω), ускорение (α) и рывок (j), рассчитанные аналогично подходу Гамбоа [13].

Угол между касательной траектории и осью *x* является арктангенсом сегмента в момент времени и принадлежит ($-\pi$, π)

$$\theta_i = \operatorname{atan2}\left(\frac{\partial x_i}{\partial y_i}\right), i = 2,\ldots,n, \qquad (1)$$

где $\partial x_i = x_i - x_{i-1}$, $\partial y_i = y_i - y_{i-1}$, $i = 2,\ldots,n$, n – количество событий в треке; $\theta_1 = 0.$

Остальные параметры рассчитываются по формулам:

$$v_{x_i} = \frac{\partial x_i}{\partial t_i}, \qquad (2)$$

$$v_{x_i} = \frac{\partial y_i}{\partial t_i}, \qquad (3)$$

$$v_i = \sqrt{{v_{x_i}}^2 + {v_{y_i}}^2}, \qquad (4)$$

$$\omega_i = \frac{\partial \theta_i}{\partial t_i}, \qquad (5)$$

$$\alpha_i = \frac{\partial v_i}{\partial t_i}, \qquad (6)$$

$$j_i = \frac{\partial \alpha_i}{\partial t_i}, \qquad (7)$$

где $\partial t_i = t_i - t_{i-1}$, $\partial \theta_i = \theta_i - \theta_{i-1}$, $\partial v_i = v_i - v_{i-1}$, $\partial \alpha_i = \alpha_i - \alpha_{i-1}$, $i = 2,\ldots,n$; n – количество событий в треке; $t_1 = 0$, $v_{x_1} = 0$, $v_{y_1} = 0$, $v_1 = 0$, $\omega_1 = 0$, $\alpha_1 = 0$, $j_1 = 0.$

Для каждого из параметров сессии v_x, v_y, v, α, j, ω, θ_i посчитано среднее, минимальное и максимальное значения, стандартное отклонение.

Кривизна между двумя соседними точками определяется как отношение изменения угла наклона и пройденного расстояния

$$c_i = \frac{\partial\theta_i}{\partial s_i}, \quad (8)$$

где $s_i = \sum_{m=1}^{i} \sqrt{\partial x_m{}^2 + \partial y_m{}^2}$; $\partial s_i = s_i - s_{i-1}$; $i = 2,\ldots,n$;

n – количество событий в треке; $s_1 = 0$.

Набор базовых характеристик включает в себя в общей сложности 50 значений.

На основании полученного вектора были рассчитаны добавочные характеристики: для каждой базовой характеристики были добавлены квантили и интерквантильный размах, доля точек, являющихся выбросами, количество изменений направления трека, энтропия и её отклонение от максимальной для ∂x_i и ∂y_i, значения суммы всех перемещений по O_x и O_y, доля точек в треке, имеющих одинаковое значение по O_x и O_y, доля углов около 0, ±45, ±90, ±135, ±180°. Всего рассчитано 150 добавочных характеристик.

В общей сложности эти 200 значений составляют входной вектор, который вычисляется для каждого трека движения мыши. Определение того, осуществляется ли движение мыши человеком, основывается на признаках, описанных выше.

Наборы данных

Для обучения и оценки разрабатываемого подхода к обнаружению ботов нужен набор данных движений мыши пользователя. Существует ограниченный перечень датасетов, включающих в себя такие данные [14], но большинство из этих наборов являются закрытыми. Проанализировав большой перечень работ и ресурсов, готовых размеченных датасетов с ботами не было найдено. Эта проблема также озвучена в работах [15–17]. Поэтому для обучения системы было решено использовать известные датасеты BALABIT и Bogazici, включающие в себя размеченные данные реальных пользователей, а также созданный самостоятельно датасет BOT1, включающие в себя данные ботов. Для тестирования используются вышеназванные датасеты и датасет UB.

Набор данных Balabit опубликован в 2016 г. [18] и относится к категории наборов неуправляемой среды, когда пользователи самостоятельно передвигают мышью без заранее полученного задания. Данный датасет включает в себя информацию о положении курсора и времени трека для 10 пользователей, работающих через клиентов удаленного рабочего стола, подключенных к серверам.

Каждое записанное событие мыши содержит шесть полей: (r_{time}, c_{time}, кнопка, состояние, x, y); r_{time} – время в секундах с момента начала сеанса, зафиксированное устройством мониторинга сети; c_{time} – это также прошедшее время, но зафиксированное клиентским устройством. Поле кнопки отображает текущее состояние кнопок мыши (левая, правая, нет кнопки), поле состояние содержит дополнительную информацию о кнопке (нажатие, отпускание, перетаскивание, движение). Поля x и y представляют собой координаты курсора на экране по осям x и y соответственно.

В табл. 1 приведен пример трека пользователя 16 наборов данных Balabit. В данном датасете представлены данные для обучения и тестирования; однако тестовые сеансы намного короче обучающих.

Таблица 1

Характеристики трека пользователя 16 наборов данных Balabit

r_{time}	c_{time}	Кнопка	Состояние	x	y
0	0	Left	Pressed	781	56
0,09045	0,09572	No button	Move	804	62
0,16318	0,16966	No button	Move	819	80
0,16897	0,17214	No button	Move	826	87
0,20479	0,20936	No button	Move	842	99
0,20479	0,20936	Left	Released	842	99
0,28075	0,29974	Left	Pressed	840	99
0,30012	0,30201	Left	Released	842	99

В табл. 2 приведен пример трека набора данных Bogazici. Для каждого пользователя есть собственный каталог для файлов сеансов, включающий в себя три каталога – обучающие данные, данные для внешнего и внутреннего тестирования.

Поведение пользователей разделено на пять основных действий: перемещение мыши (MM), перетаскивание (DD), одиночный щелчок левой кнопкой мыши (LC), одиночный щелчок правой кнопкой мыши (RC), двойной щелчок (DC).

Таблица 2

Характеристики трека набора данных Bogazici

Тип действия	Время	x	y	Кнопка	Состояние	Приложение
Перемещение	0	943	268	–	Перемещение	MS Word
Перемещение	0,107	992	175	–	Перемещение	MS Word
Нажатие	0,213	955	152	Правая	Нажатие	MS Word
Нажатие	0,298	955	152	Правая	Отпускание	MS Word
Перемещение	0,366	945	137	–	Перемещение	MS Word
Нажатие	0,428	912	129	Левая	Нажатие	MS Word

Датасет BOT1 представляет собой набор данных перемещения ботов, состоящий из 23 сессий различной длины.

Сбор данных включал в себя два этапа. Первый из них – получение данных с десктопных приложений. Второй – извлечение данных с веб-приложений. Датасет был создан с помощью различных параметризованных версий кривых Безье, сверточного автокодировщика и генеративно-состязательной сети, специализирующейся на временных рядах (TimeGAN), с использованием инструментов Selenium, Katalon,

LambdaTest, дополненных адаптированным алгоритмом WindMouse.

Данный датасет включает в себя следующие данные: тип действия (перемещение или нажатие), координаты курсора мыши по осям *x* и *y*, время, разрешение экрана, идентификаторы пользователя и сессии. Структура датасета BOT1 и пример полученного трека представлены в табл. 3. Примеры треков, созданных ботами, представлены на рис. 2.

Таблица 3

Структура датасета BOT1 и пример полученного трека

Наименование поля	Тип данных	Комментарий	Характеристики трека BOT1
actionType	Enum8	Тип действия компьютерной мыши: перемещение мыши, нажатие левой/средней (колеса)/правой/боковой кнопки, отпускание левой/средней (колеса)/правой/боковой кнопки, вращение кнопки колеса. ('MOUSE_ACTION_UNKNOWN' = –1, 'MOUSE_ACTION_MOUSEMOVE' = 0, 'MOUSE_ACTION_LBUTTONDOWN' = 1, 'MOUSE_ACTION_LBUTTONUP' = 2, 'MOUSE_ACTION_MOUSEWHEEL' = 3, 'MOUSE_ACTION_RBUTTONDOWN' = 4, 'MOUSE_ACTION_RBUTTONUP' = 5, 'MOUSE_ACTION_MBUTTONDOWN' = 6, 'MOUSE_ACTION_MBUTTONUP' = 7, 'MOUSE_ACTION_XBUTTONDOWN' = 8, 'MOUSE_ACTION_XBUTTONUP' = 9)	MOUSE_ACTION_ MOUSEMOVE
actionParamX	Int32	Координата *x* в пикселях	807
actionParamY	Int32	Координата *y* в пикселях	789
actionTime	DateTime64(9)	Временная метка совершения действия	2024-08-14 11:50:26.127972000
screenScale	Float32	Масштаб отображения экрана в процентах	100.0
screenResolutionX	Int16	Разрешение экрана по горизонтали	2560
screenResolutionY	Int16	Разрешение экрана по вертикали	1400
userId	String	Идентификатор пользователя	Cross
Session ID	String	Идентификатор сессии	sxPGZ12YFkuY8g7ihmfEKQ
ActionDeviceType	Enum8	Тип устройства, совершающий действие	MOUSE

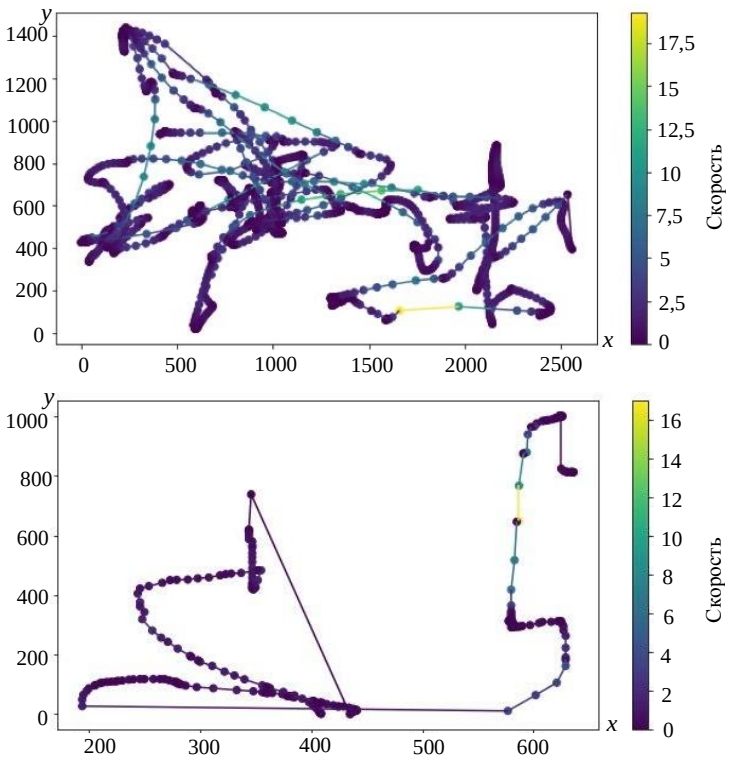


Рис. 2. Примеры движения курсора мыши ботов

Для упрощения обработки вышеназванных наборов данных необходимо уменьшить их размерности данных. По мере увеличения объема и сложности данных становится все труднее извлекать из них полезную информацию, визуализация данных становится более затратной. Методы уменьшения размерности данных решают эту проблему, позволяя представить данные в меньшем количестве измерений, оставляя наиболее важную информацию. То есть методы снижения размерности данных помогают упростить и сжать большие и сложные наборы данных, сохраняя при этом ключевую информацию. Это позволяет более эффективно работать с данными, анализировать их и визуализировать.

Для визуализации использованных наборов данных применяется метод главных компонент (PCA, principal component analysis) [20]. Основная идея PCA заключается в поиске новых признаков, называемых главными компонентами, которые имеют максимальную корреляцию с исходными данными, оставаясь при этом ортогональными друг другу. Эти главные компоненты создают новый базис в пространстве признаков, устраняя избыточную информацию и уменьшая размерность данных. Снижение размерности позволяет представить данные в двухмерном или трехмерном пространстве, что упрощает их визуальный анализ и исследование.

Согласно получившейся проекции (рис. 3), можно сделать вывод, что разработанные боты успешно имитируют поведение мыши реальными пользователями, так как все наборы данных расположены в непосредственной близости друг к другу и пересекаются.

Распределение различных признаков в применяемых наборах данных показано на рис. 4. Представленные гистограммы имеют схожий характер распределения параметров как у реальных пользователей, так и у ботов.

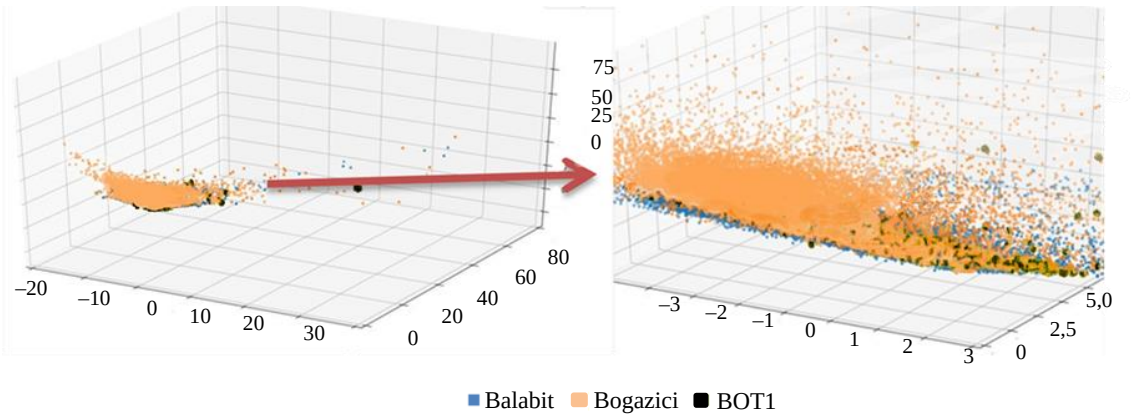
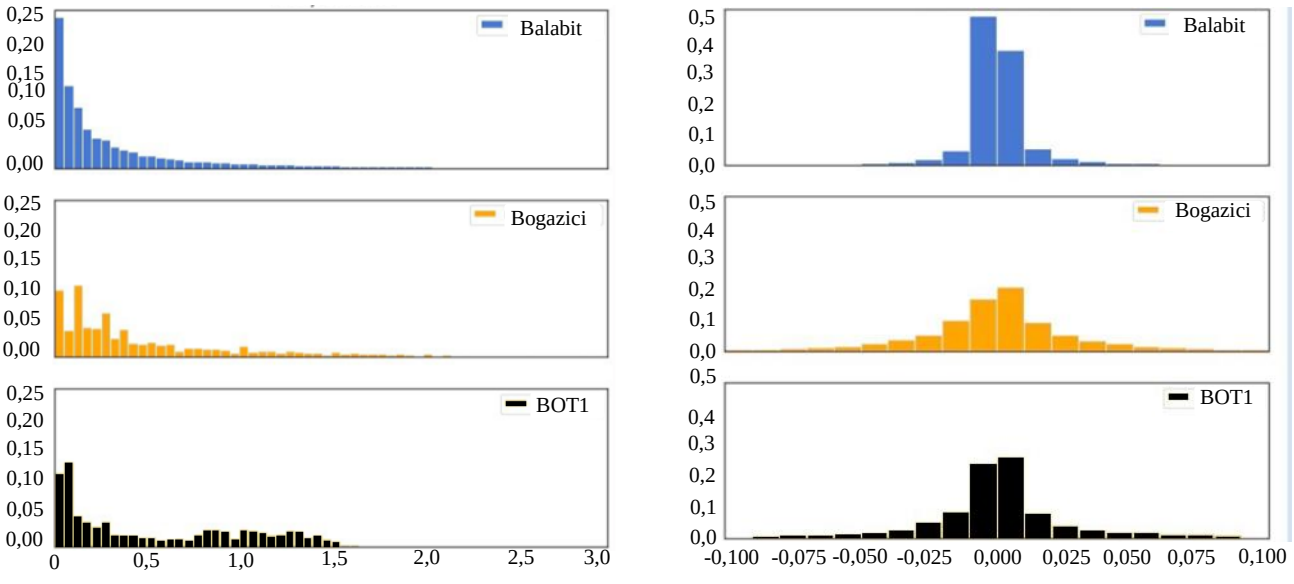


Рис. 3. Визуализация датасетов Balabit, Bogazici, BOT1, используя метод главных компонент



а *б*

Рис. 4. Распределение различных признаков в датасетах Balabit, Bogazici, BOT1: *а* – скорости, *б* – ускорения

Гистограммы распределения скорости показывают, что пользователи предпочитают перемещения курсора мыши с низкой скоростью. Причем более высокие скорости у ботов встречаются чаще, чем у реальных пользователей. Гистограммы ускорения показывают тенденцию пользователей совершать достаточно плавные движения без резких скачков скорости. Распределение рывков, т.е. скорости изменения ускорения между соседними точками, свидетельствует о преобладании небольших рывков в движении пользователей и ботов.

Заключение

Широкое распространение ботов представляет существенные риски для конфиденциальных данных, конкурентоспособности различных компаний и сервисов, формирования общественного мнения и экономической среды. Требуются комплексные меры противодействия ботам на межотраслевом и междисциплинарном уровнях.

Динамику мыши можно учитывать при детектировании ботов наряду с другими подходами, например, такими, как CAPTCHA и проверка HTTP-заголовков.

Тем не менее необходимо иметь достаточно тестовых данных, чтобы выполнить обнаружение на основании динамики мыши с хорошей точностью. Именно поэтому были выбраны общедоступные датасеты Balabit и Bogazici, а также создан собственный набор данных, описывающий перемещение мыши ботами BOT1 и включающий в себя 23 сессии.

Для дальнейших исследований было необходимо спроектировать признаковое пространство. Для этого использовались уже известные параметры в качестве базовых характеристик и в дополнение к ним предложены добавочные характеристики.

На основании этих характеристик будет приниматься решение о принадлежности сессии боту или реальному пользователю. Для этого будут применены проверка статистических гипотез и классификация данных. Также планируется использование кластеризации данных для определения принадлежности трека боту.

Работа выполнена при финансовой поддержке Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации (Минцифры России), соглашение № 40469-07/23-К от 30.06.2023.

Литература

1. Dunham K. Malicious bots: an inside look into the cyber-criminal underground of the internet / K. Dunham, J. Melnick. – Boca Raton: CRC Press, 2008. – 168 p.

2. Suchacka G. Identifying legitimate Web users and bots with different traffic profiles – an Information Bottleneck approach / G. Suchacka, J. Iwański // Knowledge-Based Systems. – 2020. – No. 197. – URL: https://doi.org/10.1016/j.knosys.2020.105875 (дата обращения: 10.06.2024).

3. Что такое боты – определение и описание [Электронный ресурс]. – Режим доступа: https://www.kaspersky.ru/resource-center/definitions/what-are-bots, свободный (дата обращения: 02.07.2024).

4. Xiao G. Bad bots: regulating the scraping of public personal information. – Cambridge: Harv. JL & Tech, 2020. – 701 p.

5. Bondy M. Bad Bots // The Project on International Peace and Security // Institute for the Theory and Practice of International Relations. – 2017. – No. 1. – P. 2016–2017.

6. Логинова А.О. Анализ существующих подходов к классификации и типологии ботов // Инновационные технологии: теория, инструменты, практика. – 2020. – Т. 1. – С. 462–467.

7. Логинова А.О. Определение атрибутов событий информационной безопасности, связанных с активностью интернет-ботов // Электронные системы и технологии: сборник матер. 58-й науч. конф. аспирантов, магистрантов и студентов БГУИР. – Минск: БГУИР, 2022. – С. 105–110.

8. Geer D. Malicious bots threaten network security // Computer. – 2005. – Vol. 38, No. 1. – P. 18–20.

9. Kolomeets M. Analysis of the malicious bots market / M. Kolomeets, A. Chechulin // IEEE 29th conference of open innovations association (FRUCT). – 2021. – P. 199–205. – URL: https://www.researchgate.net/publication/351855704_Analysis_of_the_Malicious_Bots_Market, свободный (дата обращения: 13.06.2024).

10. 2023 Imperva Bad Bot Report [Электронный ресурс]. – Режим доступа: https://www.imperva.com/resources/resource-library/reports/2023-imperva-bad-bot-report/, свободный (дата обращения: 12.06.2024).

11. OWASP Automated Threats to Web Applications [Электронный ресурс]. – Режим доступа: https://owasp.org/www-project-automated-threats-to-web-applications/, свободный (дата обращения: 20.06.2024).

12. Efficient on-the-fly Web bot detection / G. Suchacka, A. Cabri, S. Rovetta, F. Masulli // Knowledge-Based Systems. – 2021. – Vol. 223. – URL: https://www.sciencedirect.com/science/article/pii/S0950705121003373, свободный (дата обращения: 20.06.2024).

13. Gamboa H. A behavioral biometric system based on human-computer interaction / H. Gamboa, A. Fred // Biometric Technology for Human Identification. – 2004. – Vol. 5404. – P. 381–392.

14. Afanaseva N.S. Bot Detection Using Mouse Movements / N.S. Afanaseva. P.S. Lozhnikov // 2023 Dynamics of Systems, Mechanisms and Machines (Dynamics). – IEEE, 2023. – P. 1–4.

15. Detecting Web Bots via Mouse Dynamics and Communication Metadata / A. See, T. Wingarz, M. Radloff, M. Fischer // IFIP International Conference on ICT Systems Security and Privacy Protection. – 2023. – P. 73–86.

16. Antal M. Intrusion detection using mouse dynamics / M. Antal, E. Egyed-Zsigmond // IET Biometrics. – 2019. – Vol. 8, No. 5. – P. 285–294.

17. Mouse dynamics behavioral biometrics: A survey / S. Khan, C. Devlen, M. Manno, D. Hou // ACM Computing Surveys. – 2024. – Vol. 56, No. 6. – P. 1–33.

18. Antal M. Mouse dynamics based user recognition using deep learning / M. Antal, N. Fejér // Acta Universitatis Sapientiae, Informatica. – 2020. – Vol. 12, No. 1. – P. 39–50.

19. Kılıç A.A. Bogazici mouse dynamics dataset / A.A. Kılıç, M. Yıldırım, E. Anarım // Data in Brief. – 2021. – Vol. 36. – P. 107094.

20. Maćkiewicz A. Principal components analysis (PCA) / A. Maćkiewicz, R. Waldemar // Computers & Geosciences. – 1993. – Vol. 19, No. 3. – P. 303–342.

Афанасьева Наталья Сергеевна
Ст. преп. кафедры информационной безопасности
Омского гос. ун-та путей сообщения (ОмГУПС)
Маркса пр-т, 35, г. Омск, Россия, 644046
ORCID: 0009-0003-7591-486X
Тел.: +7-904-324-06-66
Эл. почта: nati_dik@mail.ru

Ложников Павел Сергеевич
Д-р техн. наук, проф., зав. каф. комплексной защиты информации ОмГТУ
Мира пр-т, 11, г. Омск, Россия, 644050
ORCID: 0000-0001-7878-1976
Тел.: +7-913-605-33-33
Эл. почта: lozhnikov@mail.ru

Afanaseva N.S., Lozhnikov P.S.
Datasets for bot detection using mouse behavior

Bot detection based on the dynamic characteristics of computer mouse cursor movement is considered. The basic dynamic characteristics and 150 additional ones obtained on their basis are presented. Publicly available datasets used to detect bots based on mouse cursor movement characteristics are studied. Self-created BOT1 and BOT2 datasets used for system training are described. To confirm the applicability of these data sets, the dimensionality reduction and the distribution comparison of key dynamic characteristics of the sets are performed.
Keywords: bot; malicious bot; mouse movement; track, mouse dynamics datasets.
DOI: 10.21293/1818-0442-2024-27-3-118-124

References

1. Dunham K., Melnick, J. *Malicious bots: an inside look into the cyber-criminal underground of the internet*. Boca Raton, CRC Press, 2008, 168 p.

2. Suchacka G., Iwański J. Identifying legitimate Web users and bots with different traffic profiles – an Information Bottleneck approach. *Knowledge-Based Systems,* 2020, no. 197. Available at: https://doi.org/10.1016/j.knosys.2020.105875, free (Accessed: June 10, 2024).

3. What are bots. Available at: https://www.kaspersky.ru/resource-center/definitions/what-are-bots, free (Accessed 2 July 2024).

4. Xiao G. *Bad Bots: Regulating the Scraping of Public Personal Information*. Cambridge, Harv. JL & Tech, 2020, 701 p.

5. Bondy M. Bad Bots. The Project on International Peace and Security, Institute for the. *Theory and Practice of International Relations*, 2017, pp. 2016–2017.

6. Loginova A.O. *Analiz sushchestvuyushchih podhodov k klassifikacii i tipologii botov* [The analysis of existing approaches to bots classification and typology]. *Innovative Technologies: Theory, Tools, Practice,* 2020, vol. 1, pp. 462–467 (in Russ.).

7. Loginova A.O. *Opredelenie atributov sobytij informacionnoj bezopasnosti svyazannyh s aktivnostyu internet botov* [Defining information security events attributes related to the activity of internet bots]. *Electronic Systems and Technologies. Proceedings of the 58th Scientific Conference of Postgraduate, Graduate and Undergraduate Students of Belarusian State University of Informatics and Radioelectronics*, 2022, pp. 105–110 (in Russ.).

8. Geer D. Malicious bots threaten network security. *Computer*, 2005, vol. 38, no. 1, pp. 18–20.

9. Kolomeets M., Chechulin A. Analysis of the malicious bots market. *IEEE 29th Conference of Open Innovations Association (FRUCT)*, 2021, pp. 199–205. Available at: https://www.researchgate.net/publication/351855704_Analysis_of_the_Malicious_Bots_Market, free (Accessed: June 13, 2024).

10. 2023 Imperva Bad Bot Report. Available at: https://www.imperva.com/resources/resource-library/reports/2023-imperva-bad-bot-report/, free (Accessed: June 12, 2024).

11. OWASP Automated Threats to Web Applications. Available at: https://owasp.org/www-project-automated-threats-to-web-applications/, free (Accessed: June 20, 2024).

12. Suchacka G., Cabri A., Rovetta S., Masulli F. Efficient on-the-fly Web bot detection. *Knowledge-Based Systems*, 2021, vol. 223. Available at: https://www.sciencedirect.com/science/article/pii/S0950705121003373, free (Accessed: June 20, 2024).

13. Gamboa H., Fred A. A behavioral biometric system based on human-computer interaction. *Biometric Technology for Human Identification*, SPIE, 2004, vol. 5404, pp. 381–392.

14. Afanaseva N.S., Lozhnikov P.S. Bot Detection Using Mouse Movements. *2023 Dynamics of Systems, Mechanisms and Machines (Dynamics)*, IEEE, 2023, pp. 1–4.

15. See A., Wingarz T., Radloff M., Fischer M. Detecting Web Bots via Mouse Dynamics and Communication Metadata. *IFIP International Conference on ICT Systems Security and Privacy Protection*, 2023, pp. 73–86.

16. Antal M., Egyed-Zsigmond E. Intrusion detection using mouse dynamics. *IET Biometrics,* 2019, vol. 8, no. 5, pp. 285–294.

17. Khan S., Devlen C., Manno M., Hou D. Mouse dynamics behavioral biometrics: A survey. *ACM Computing Surveys*, 2024, vol. 56, no. 6, pp. 1–33.

18. Antal M., Fejér N. Mouse dynamics based user recognition using deep learning. *Acta Universitatis Sapientiae, Informatica*, 2020, vol. 12, no. 1, pp. 39–50.

19. Kılıç A.A., Yıldırım M., Anarım E. Bogazici mouse dynamics dataset. *Data in Brief*, 2021, vol. 36, p. 107094.

20. Maćkiewicz, A., Waldemar R. Principal components analysis (PCA). *Computers & Geosciences*, 1993, vol. 19, no. 3, pp. 303–342.

Natalia S. Afanaseva
Lecturer, Department Information Security
Omsk State Transport University
35, Marx st., Omsk, Russian Federation, 644046
ORCID: 0009-0003-7591-486X
Phone: +7-904-324-06-66
Email: nati_dik@mail.ru

Pavel S. Lozhnikov
Doctor of Science in Engineering, Professor,
Department of Comprehensive Information Protection
11, Mira st., Omsk, Russian Federation, 644050
ORCID: 0000-0001-7878-1976
Phone: +7-913-605-33-33
Email: lozhnikov@mail.ru