

АНАЛІЗ МЕТОДІВ ПРОТИДІЇ АВТОМАТИЗОВАНОМУ СКАНУВАННЮ ВЕБ-САЙТІВ

Шипунов М.Ю.

Національний аерокосмічний університет ім. М. Е. Жуковського «ХАІ»
Науковий керівник Цуранов М.В.

Актуальність. З кожним роком кількість відвідувачів Інтернет ресурсів невинно зростає. Однією з причин є доступність інформаційних технологій звичайним користувачам та зниження плати за доступ до мережі Інтернет. Згідно даних аналітичної компанії Similarweb, у вересні 2022 року найпопулярнішими сайтами серед українців є Google, YouTube, Facebook, Ukr.net, Sinoptik, OLX, Deepstatemap та Wikipedia [1]. Згідно цих даних видно, що серед користувачів популярністю користується розважальний контент, портали новин, електронна пошта, соціальні мережі та торговельні майданчики. Збільшення кількості користувачів сайту призводить до появи автоматизованих систем, які мають на меті злочинні наміри. Такі боти можуть мати різні цілі, а саме: збирання інформації про товари (для перепродажу їх по більшій ціні), сканування вразливостей сайту (не завжди з метою нанесення збитків, також для перевірки захищеності) або генерація спам повідомлень (поширення рекламної інформації у коментарях сайту). Виходячи з цього стає зрозумілим, що сайти повинні мати не тільки потужні сервери для безперервної роботи, а також механізми, які здатні протидіяти системам автоматизованого сканування.

Мета роботи полягає у аналізі методів протидії автоматизованому скануванню веб-сайтів.

Основні положення. Автоматизовані системи аналізу сайтів поділяються на 2 категорії: ті, які допомагають власнику сайту та ті, які наносять шкоду. До корисних відносяться індексатори пошукових систем, які аналізують веб-ресурси з метою додавання їх до результатів пошуку користувачів [2]. До шкідливих відносяться боти, метою яких можуть бути: поширення реклами у коментарях, сканування товарів з метою продажу їх по вищій ціні або сканування вразливостей сайту. Важливо зазначити, що не існує чітких меж між корисними та шкідливими ботами. Таким чином пентестингові боти можуть одночасно відноситися до обидвох категорій, в залежності від того, хто їх використовує (власник сайту для перевірки рівня захищеності або зловмисник з метою нанесення шкоди). Окрім прямого нанесення збитків веб-ресурсу, боти також створюють зайве навантаження на сервери. Це може призвести як до уповільнення завантаження сайту, так і до повної відсутності доступу до нього реальними користувачами.

Згідно даних компанії Imperva, яка займається розробкою корпоративного програмного забезпечення у сфері інформаційної безпеки,

кількість мережевого трафіку у 2021 році, який припадає на мережевих ботів складає 42,6%, що на 1,5% більше порівняно з 2020 роком [3]. Окрім цього, найпопулярнішими сферами застосування Інтернет ботів є спортивні сайти, сайти з азартними іграми та ресурси туристичних агенств.

На сьогоднішній день, для захисту веб-сайту від мережевих ботів існує багато методів та готових рішень. Серед них варто виділити: чорні списки (містять в собі IP адреси хостів, які треба заблокувати), сигнатурне блокування (блокування ботів на основі сигнатур у HTTP запитах), використання WAF (сукупність моніторів та фільтрів, призначених для виявлення та блокування мережевих атак на прикладному рівні) та використання технології Captcha (автоматизований тест Тюрінга, який має на меті розрізнення комп'ютерів та людей) [4].

Висновки. Під час аналізу даних компанії Imperva було виявлено, що найчастіше системи автоматизованого сканування мають на меті заробіток для їх власників. Таким чином, наприклад, автоматичне сканування облікових записів сайту онлайн казино дозволяє викрасти гроші користувачів шляхом перехоплення даних акаунту. Окрім цього, виходячи з розглянутої статистики, з кожним роком кількість трафіку, який генерується шкідливими мережевими ботами лише збільшується. Але, нажаль не існує єдиного та універсального рішення для боротьби з автоматизованими системами. Це зумовлено тим, що під час вибору методів захисту треба враховувати тип мережевих ботів, яким треба протидіяти. Таким чином для боротьби з ботами, які маскуються під користувачів треба використовувати Captcha, проти спамерів – чорні списки IP адрес, а для боротьби з сканерами вразливостей доцільніше використовувати WAF.

Список літератури

1. Рейтинг топ-сайтів. URL: <https://www.similarweb.com/ru/top-websites/ukraine/> (дата звернення: 19.10.2022);
2. Google Indexing API. URL: <https://developers.google.com/search/apis/indexing-api/v3/quickstart> (дата звернення: 20.10.2022);
3. Bad Bot Report 2021: The Pandemic of the Internet. URL – <https://www.imperva.com/blog/bad-bot-report-2021-the-pandemic-of-the-internet/> (дата звернення: 20.10.2022);
4. V. Pevnev, O. Popovichenko, and Y. Tsokota, “WEB APPLICATION PROTECTION TECHNOLOGIES”, A.I.S., vol. 4, no. 1, pp. 119–123, Mar. 2020.

Відомості про авторів

Шипунов Микита Юрійович, магістрант кафедри комп'ютерних систем, мереж і кібербезпеки, м.т. 0980344388, m.shypunov@student.csn.khai.edu
 Цуранов Михайло Віталійович, дослідник у сфері ІБ