УДК 004.056
ГЕЛЬДТ Станіслав Володимирович,
курсант 2 курсу факультету № 4
Харківського національного університету внутрішніх справ;
ОНИЩЕНКО Юрій Миколайович,
кандидат наук з державного управління, доцент,
доцент кафедри кібербезпеки та DATA-технологій факультету № 6
Харківського національного університету внутрішніх справ
https://orcid.org/0000-0002-7755-3071;

АНАЛІЗ ЛОГ-ФАЙЛІВ ДЛЯ ПРОТИДІЇ КІБЕРАТАКАМ

Згідно з даними Державної служби спеціального зв'язку та захисту інформації, за період війни вже сталося майже втричі більше різного роду хакерських атак, ніж за аналогічний період минулого року [1]. Це вплинуло на роботу таких сайтів та порталів, як: Приват24, Ощадбанк, Дія, Міністерство оборони України, ЗСУ [2]. Мета скоординованих нападів – порушення роботи державних та військових органів, підрив довіри населення до державних інституцій. Таким чином, влада понесла колосальні збитки у вигляді злитих конфіденціальних даних багатьох українців.

Головною задачею цифрової країни є безпека приватної інформації, тому найбільш важливо вміти якомога швидше реагувати на кіберінциденти. Одним із найкращих методів виявлення слідів втручання у систему є перевірка та аналіз певних файлів.

Лог-файл – це створений комп'ютером файл даних, який містить інформацію про моделі використання, дії та операції в програмі, сервері чи іншому пристрої [3]. Проаналізувавши сучасні операційні системи, можна прийти до висновку, що операційна система Linux ефективно налаштована на створення та зберігання файлів журналів. Linux створює безперервну часову шкалу подій, які відбуваються в системі, включаючи кожну подію, пов'язану з сервером, ядром і запущеними програмами. Linux поділяє події на чотири різні категорії:

- журнали додатків;
- журнали подій;
- журнали обслуговування;
- системні журнали.

Щоб зрозуміти важливість інформації та способи її обробки, пропонується розібрати один з лог-файлів сервера Apache. Apache – відкритий вебсервер Інтернет для UNIX-подібних, Microsoft Windows, Novell NetWare та інших операційних систем [4]. Файл знаходиться за абсолютним шляхом – /var/log/apache2/access.log. Завжди має певну структуру, а саме:

– IP-адреса хоста, який запросив сторінку;
– дата, час та GMT;
– сторінка, яку запросив хост;
– версія протоколу;
– код стану;
– розмір файлу в байтах;
– сторінка, з якої посилається хост;
– агент користувача, ідентифікований браузером [5].

Завдяки структурі, файл можна аналізувати за допомогою утиліти cut. Наприклад, команда нижче виведе лише IP-адреси з лог-файлу:

cut –d' ' –f 1 /var/log/apache2/access.log

Параметр –d вказується для визначення роздільника тексту, -f для номера стовпця.

Для виводу інформації щодо вебсторінок, які відвідувала певна IP-адреса, використовується утиліта awk:

awk '$1 == "192.168.0.3" {print $0}' /var/log/apache2/access.log | cut –d' ' –f 7

Awk – це мова програмування, призначена для сканування та обробки зразків [6]. В команді вище '192.168.0.3' – тестова IP-адреса, яку ми маємо проаналізувати.

Якщо хост зайшов на кожну сторінку лише по 1 разу, це може вказувати на вебсканер або клонування сайту. Website Crawler – програма, що є складовою частиною пошукової системи та призначена для обходу сторінок інтернету з метою занесення інформації про них до бази даних.

awk '$9 == "404" {print $0}' /var/log/apache2/access.log | cut –d' ' –f 1

За допомогою цієї команди можна проаналізувати IP-адреси, яким повернувся статус 404 (Not Found). Якщо хост виводиться забагато разів, є сенс перевірити які саме сторінки він відвідував. Web Directory Enumeration – атака методом застосування грубої сили на приховані файли та каталоги шляхом послідовного відвідування сторінок, визначених у списку спеціального словника [7].

awk '$1 == "192.168.0.3" {print $0}' /var/log/apache2/access.log | cut –d' ' –f 12

Команда, наведена вище, аналізує агента користувача, ідентифікованого браузером. Таким чином, можна побачити нестандартні браузери та боти, які не являються прямим користувачами.

tail –f /var/log/apache2/access.log | egrep –line-buffered 'HTTP/. 404' | cut –d' ' –f4-7*

Утиліта tail з параметром –f аналізує лог-файл в режимі реального часу. Утиліта egrep знаходить збіг з рядком *'HTTP/.* 404'* та виводить в термінал дату, час, тип запиту та сторінку.

Файли журналу та Системні журнали є невід'ємною частиною захисту системи на основі перегляду та аналізу. Лише однією командою, адміністратор може контролювати кожен запит користувача та швидко реагувати на певні аномалії. Завдяки автоматизації процесів за допомогою утиліти mail, нестандартні строки можуть бути відправлені на електрону пошту та проаналізовані в короткий період часу.

Список використаних джерел

1. За час війни кількість хакерських атак в Україні зросла втричі // Економічна правда : вебсайт. URL: https://www.epravda.com.ua/news/2022/04/3/685157/ (дата звернення: 21.04.2022).

2. Масована DDoS-атака на ПриватБанк та Ощадбанк. Сайт Міноборони не відкривається // Ліга.Tech : вебсайт. URL: https://tech.liga.net/ua/ukraine/article/massirovannaya-ddos-ataka-na-privatbank-i-oschadbank-sayt-minoborony-ne-otkryvaetsya (дата звернення: 21.04.2022).

3. DevOps and Security Glossary Terms // Log File : вебсайт. URL: https://www.sumologic.com/glossary/log-file/ (дата звернення: 21.04.2022).

4. Apache HTTP Server // Wikipedia : вебсайт. URL: https://uk.wikipedia.org/wiki/Apache_HTTP_Server (дата звернення: 21.04.2022).

5. Bash-скрипты // Хабр : вебсайт. URL: https://habr.com/ru/company/ruvds/blog/325522/ (дата звернення: 21.04.2022).

6. AWK // Wikipedia : вебсайт. URL: https://uk.wikipedia.org/wiki/AWK (дата звернення: 21.04.2022).

7. Web Directory Enumeration // Gitbooks : вебсайт. URL: https://www.epravda.com.ua/news/2022/04/3/685157/ (дата звернення: 21.04.2022).

Одержано 01.05.2022