УДК 004.7

Гнатюк В.О., магістрант
Державний університет «Житомирська політехніка»

СИСТЕМА АНАЛІЗУ МЕРЕЖЕВОГО ТРАФІКУ З ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ ЗАДАЧ SIGINT

У сучасних умовах зростання кількості кіберзагроз особливого значення набувають ефективні системи аналізу мережевого трафіку. Традиційні підходи, що базуються на сигнатурному аналізі, дозволяють виявляти лише відомі типи атак і є малоефективними проти нових або модифікованих загроз. Це обумовлює необхідність використання інтелектуальних методів аналізу, зокрема машинного навчання, у поєднанні з підходами Signals Intelligence (SIGINT)[2].

Метою роботи є розробка системи аналізу мережевого трафіку, яка використовує алгоритми машинного навчання для виявлення аномалій та класифікації мережевих подій.

Запропонована система передбачає кілька основних етапів: збір даних, попередню обробку, виділення ознак та класифікацію. Збір мережевого трафіку може здійснюватися за допомогою аналізатора мережевих протоколів Wireshark із подальшим збереженням даних у форматі PCAP. Це дозволяє отримати детальну інформацію про мережеві пакети для подальшого аналізу[1].

На етапі попередньої обробки виконується фільтрація, нормалізація та підготовка даних. Основними характеристиками, що використовуються для аналізу, є IP-адреси джерела і призначення, тип протоколу, розмір пакета, інтенсивність передачі та часові інтервали між пакетами. Ці параметри формують набір ознак для моделей машинного навчання.

Для класифікації мережевого трафіку застосовуються алгоритми машинного навчання, зокрема дерева рішень, випадковий ліс та метод опорних векторів. Навчання моделей здійснюється на основі підготовлених наборів даних, що містять приклади як нормального, так і аномального трафіку. У процесі навчання модель формує правила, які дозволяють відрізняти типи мережевої активності[3].

Після навчання система здатна аналізувати нові дані та визначати їх належність до певного класу: нормальний трафік, сканування мережі або потенційна атака. Аналіз може виконуватися як у реальному часі, так і в режимі офлайн для попередньо збережених даних. У разі виявлення відхилень система формує попередження для подальшого реагування.

Для тестування системи може бути створене лабораторне середовище із використанням засобів віртуалізації або мережевих емуляторів, таких як GNS3. Це дозволяє моделювати різні сценарії мережевої активності та формувати навчальні вибірки.

Запропонований підхід має низку переваг. По-перше, система здатна виявляти нові типи атак без наявності попередніх сигнатур. По-друге, використання машинного навчання дозволяє ефективно обробляти великі обсяги даних та знаходити складні закономірності. По-третє, система може бути реалізована без використання спеціалізованого обладнання.

Отже, поєднання методів машинного навчання з підходами SIGINT дозволяє створити ефективну систему аналізу мережевого трафіку. Подальший розвиток таких систем пов’язаний із вдосконаленням моделей, розширенням набору ознак та інтеграцією з іншими засобами кіберзахисту.

Список використаних джерел

1. Network Security Through Data Analysis: Building Situational Awareness / Майкл Коллінз – O’Reilly Media, 2014 – 348 с

2. Визначення SIGINT [Електронний ресурс] – Режим доступу до ресурсу: https://www.assured-systems.com/faq/what-is-signal-intelligence-sigint-and-why-does-it-matter/

3. Machine Learning and Security: Protecting Systems with Data and Algorithms / Кларенс Чіо, Девід Фрімен – O’Reilly Media, 2018 – 370 с