УДК 004.4

РОЗРОБКА ПРОГРАМИ ДЛЯ ВИЯВЛЕННЯ ШКІДЛИВОГО ОНЛАЙН ТРАФІКУ

Казимир А.А. – гр. МгКІ-24, магістрант, andrewkazimir@gmail.com
Стаценко Д.В. – к.т.н., доцент., statsenko.dv@knutd.edu.ua
Київський національний університет технологій та дизайну

Метою роботи є створення програмного засобу для онлайн-виявлення шкідливого мережевого трафіку з використанням сучасних методів аналізу мережевих даних.

Шкідливий мережевий трафік - це мережевий трафік, який спеціально розроблений для атаки на мережу, систему або додаток [1]. До основних типів шкідливого трафіку належать: атаки типу Denial of Service (DoS/DDoS), спроби сканування портів, несанкціонований доступ до систем (Brute Force, SQL-ін’єкції), експлуатація вразливостей веб-додатків, а також фішингові та ботнет-активності. Залежно від цілі зловмисника, шкідливий трафік може бути спрямований на блокування сервісів, отримання конфіденційної інформації, розповсюдження шкідливого ПЗ або приховане використання ресурсів (наприклад, майнінг криптовалюти) [2]. Класифікація такого трафіку зазвичай базується на його структурних ознаках (протокол, розмір пакета, час між запитами, напрямок потоку) або на поведінкових характеристиках - шаблонах активності користувача чи системи. Це дозволяє формувати статистичні моделі за якими можна відрізняти звичайну мережеву поведінку від потенційно шкідливої.

Серед основних підходів до виявлення шкідливого трафіку розрізняють чотири методи:

1) Сигнатурний метод – метод, який базується на пошуку відомих шаблонів (сигнатур) атак. Його перевагою є висока точність при виявленні відомих загроз, але недоліком є неможливість розпізнати нові типи атак, що не мають сигнатури.
2) Аномальний метод – метод, який використовує статистичний або машинний аналіз для виявлення відхилень від «норми». Він ефективний для нових або модифікованих атак, однак схильний до хибних спрацьовувань.
3) Поведінковий метод – метод, що орієнтується на аналіз дій користувача або пристрою, виявляючи незвичні шаблони поведінки. Такий підхід широко застосовується у сучасних системах моніторингу безпеки [3, 4].

4) Гібридний метод – метод, який поєднує переваги попередніх трьох, використовуючи сигнатури для фіксації відомих загроз і машинне навчання для пошуку аномалій.

Саме гібридний метод виявлення є найбільш перспективними для сучасних систем захисту мереж.

Системи виявлення та запобігання вторгненням (IDS/IPS) стали стандартом у сфері мережевої безпеки. Наприклад, Snort - одна з найпоширеніших IDS, що використовує сигнатурний метод [5]. Її гнучка конфігурація дозволяє аналізувати трафік у режимі реального часу, однак вона обмежена при роботі з новими загрозами. Ще одна поширена IDS – Suricata. Вона дозволяє одночасно здійснювати виявлення атак і глибокий аналіз трафіку, що робить її ефективнішою у масштабних мережах. Попри високу ефективність, традиційні IDS мають обмеження у швидкості оновлення баз та адаптації до нових форм атак [6]. Саме тому активно розвиваються рішення на основі машинного та глибинного навчання.

Методи машинного навчання дозволяють автоматично виявляти закономірності у даних та виявляти аномалії без попереднього знання сигнатур атак. Найпопулярнішими алгоритмами є Decision Tree, Random Forest, Support Vector Machine та нейронні мережі. Особливо ефективними виявилися згорткові нейронні мережі (CNN), які можуть сприймати мережевий трафік як зображення, аналізуючи «патерни» у послідовностях байтів. Це забезпечує точнішу ідентифікацію складних атак і зменшує кількість хибних спрацювань. Застосування глибинного навчання у системах мережевої безпеки дозволяє перейти від статичних сигнатур до динамічних моделей, здатних адаптуватися до нових видів загроз [7]. Згорткові нейронні мережі обрано завдяки їх здатності ефективно обробляти структуровані дані та виділяти ознаки, недоступні класичним алгоритмам. CNN здатна «бачити» мережевий трафік як двовимірне зображення, де кожен байт пакета відповідає яскравості пікселя. Це дозволяє автоматично навчити мережу розпізнавати приховані закономірності, характерні для шкідливої активності.

Архітектура моделі включає кілька згорткових шарів для вилучення ознак, шари пулінгу для зменшення розмірності, та повнозв’язний шар для прийняття рішення [8]. На виході використовується сигмоїдна функція активації, яка визначає, чи є трафік шкідливим або нормальним.

Для навчання нейронної мережі використовують відкриті набори даних KDD Cup 99 та CICIDS2017, що містять як нормальний, так і шкідливий трафік різних типів (DoS, DDoS, PortScan, Web Attack, Infiltration). Дані попередньо очищують, та нормалізовують до діапазону [0;1], після чого перетворено у

візуальний формат - зображення розміром 64на64 пікселі. Це дозволяє подати байтові послідовності у вигляді, яки є придатним для CNN.

Отже, перспективним напрямом розвитку систем захисту інформації є поєднання класичних механізмів аналізу трафіку з інтелектуальними моделями машинного навчання. Саме цей підхід покладено в основу подальшого проєктування програмного застосунку для автоматизованого виявлення шкідливого мережевого трафіку в режимі реального часу.

Список використаних джерел:

1. https://www.mdpi.com/1099-4300/25/5/821
2. https://www.researchgate.net/profile/Nitin-Choudhury/publication/375883185_Malicious_Traffic_Classification_Using_Convolutional_Neural_Network/links/6560e12fb86a1d521b0549fa/Malicious-Traffic-Classification-Using-Convolutional-Neural-Network.pdf
3. A Primer on the Signature Method in Machine Learning / I. Chevyrev, A. Kormilitzin. *arXiv*, 2016. 47 p. DOI: 10.48550/arXiv.1603.03788.
4. https://www.dnssense.com/limitations-signature-based-threat-detection
5. Chandola, V., Banerjee, A., & Kumar, V. (2009). “Anomaly Detection: A Survey.” ACM Computing Surveys. — comprehensive survey of anomaly detection methods across domains.
6. https://www.ibm.com/think/topics/intrusion-detection-system
7. Al-Jarrah, O. Machine Learning Techniques for Network Intrusion Detection / O. Al-Jarrah [та ін.]. *ResearchGate*, 2020. 15 p. URL: https://www.researchgate.net/publication/344158644_Machine_Learning_Techniques_for_Network_Intrusion_Detection
8. Mohammadpour L., Ling T. C., Liew C. S., Aryanfar A. «A Survey of CNN-Based Network Intrusion Detection». *Applied Sciences*, 2022, 12(16): 8162. DOI:10.3390/app12168162.