УДК 004.0563:004.8

Костромицький Олександр Сергійович,

здобувач другого (магістерського) рівня вищої освіти
Харківського національного університету радіоелектроніки

Піксасов Михайло Михайлович,

к.т.н., доцент кафедри кібербезпеки Харківського національного автомобільно-дорожнього університету

ОЦІНЮВАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ АНОМАЛІЙ МЕРЕЖЕВОГО ТРАФІКУ

Анотація. Досліджено можливості автоматизованого виявлення мережевих атак шляхом бінарної класифікації характеристик мережевих з'єднань із використанням алгоритмів Random Forest та багатошарового перцептрона (MLP). Розглянуто походження експериментальних даних, особливості підготовки ознак і методи оцінювання результатів класифікації. За результатами експериментальних досліджень на наборах даних NSL-KDD та CICIDS2017 встановлено високу точність виявлення мережевих атак обома алгоритмами. Обґрунтовано некоректність безпосереднього порівняння ефективності моделей, досліджених на різних наборах даних. Проаналізовано вплив хибнопозитивних і хибнонегативних рішень на результативність виявлення кіберзагроз та визначено необхідність додаткової перевірки стійкості моделей на незалежному мережевому трафіку.

Ключові слова: мережеві атаки, аналіз аномалій, Random Forest, багатошаровий перцептрон, NSL-KDD, CICIDS2017, виявлення вторгнень.

Забезпечення кіберстійкості корпоративних мереж потребує виявлення не лише відомих сигнатур атак, а й відхилень у параметрах мережевих з'єднань. Сигнатурні механізми ефективні для відомих шаблонів зловмисної активності, проте залежні від своєчасного оновлення правил. Методи машинного навчання дають змогу класифікувати потоки за сукупністю статистичних і протокольних ознак, хоча якість такого виявлення залежить від репрезентативності навчального трафіку й умов експлуатації. Для подальшого опрацювання попереджень у системах моніторингу безпеки важливо також забезпечувати кореляцію та аналіз подій безпеки [1, 2].

Метою дослідження є оцінювання можливостей Random Forest та нейронної мережі MLP для розрізнення легітимного й атакувального

мережевого трафіку та визначення умов коректної інтерпретації показників класифікації. Для цього виконано два програмні експерименти з навчанням і тестуванням моделей на спеціалізованих наборах даних.

Експериментальну основу становлять набори NSL-KDD і CICIDS2017. NSL-KDD містить характеристики мережевих з'єднань та їхні класи, успадковані від завдань виявлення вторгнень KDD, його структурованість полегшує опрацювання категоріальних і числових ознак. CICIDS2017 включає марковані потоки звичайного трафіку та відтворених сценаріїв атак, зокрема DoS/DDoS, перебору облікових даних і сканування [3, 4]. У цих наборах відмінні періоди формування, структура ознак, баланс класів і сценарії атак, тому перенесення висновків між ними потребує окремої перевірки.

Для моделі Random Forest використано записи NSL-KDD. Реалізацію виконано мовою Python із застосуванням scikit-learn. Якість моделі оцінювали за матрицею помилок на тестових записах. За результатами тестування алгоритму Random Forest встановлено високу точність класифікації нормального та шкідливого мережевого трафіку. Переважну більшість тестових записів класифіковано правильно, проте зафіксовано поодинокі хибнопозитивні спрацювання та випадки невиявлення атак. Отримані результати свідчать про високу ефективність алгоритму в досліджуваних умовах, водночас підтверджуючи, що навіть за високої точності класифікації зберігається ризик пропуску окремих кіберзагроз.

Другий експеримент проведено з використанням багатошарового перцептрона (MLP) на підготовленій вибірці CICIDS2017. Перед навчанням здійснено нормалізацію числових характеристик мережевих потоків, а початкові мітки зведено до двох класів: нормальна активність та атака. Програмну реалізацію виконано мовою Python із використанням бібліотек нейромережевого моделювання. За результатами тестування встановлено високу точність класифікації нормального та шкідливого мережевого

трафіку. Переважну більшість записів класифіковано правильно, при цьому зафіксовано незначну кількість хибнопозитивних спрацювань і випадків невиявлення атак, причому останні переважали. Отримані результати свідчать про високу результативність моделі на досліджуваній вибірці, однак через округлення вихідних показників підсумкову точність слід розглядати як наближену оцінку.

Порівняння отриманих результатів не дає підстав стверджувати про перевагу багатошарового перцептрона (MLP) над алгоритмом Random Forest, оскільки моделі тестувалися на різних наборах даних без проведення єдиного незалежного випробування. Для об’єктивного оцінювання їхньої ефективності необхідно забезпечити використання спільного набору даних, однакових умов формування навчальної та тестової вибірок, узгоджених процедур попередньої обробки, а також врахувати тривалість навчання, час класифікації та характер помилок для кожного типу атаки. Особливу увагу слід приділити запобіганню витоку інформації між навчальними та тестовими даними й перевірці здатності моделей зберігати ефективність за умов зміни характеристик мережевого трафіку.

Таким чином, отримані результати демонструють можливість побудови програмних модулів бінарної класифікації мережевої активності на основі Random Forest і MLP у контрольованих умовах. Практичне значення полягає у визначенні послідовності підготовки мережевих ознак, навчання класифікаторів та інтерпретації їхніх помилок. Для використання таких моделей у реальному моніторингу потрібні незалежна перевірка на нових потоках, калібрування порогів сповіщень і інтеграція результатів у процес аналізу інцидентів із участю фахівців безпеки [1, 2].

Список використаних джерел

1. Сєвєрінов О.В., Чубенко О.В. Аналіз проблем використання SIEM систем. Проблеми інформатизації: тези доп. восьмої міжнар. наук.-техн.

конф. 2020. Т. 1. - С. 111.
2. Sievierinov O.V., Ovcharenko M. Analysis of Correlation Rules in Security Information and Event Management Systems. Computer and Information Systems and Technologies: Proc. of the 4th Int. Sci. and Tech. Conf. Kharkiv: NURE, 2020. P. 24–25.
3. Canadian Institute for Cybersecurity. NSL-KDD Dataset. University of New Brunswick. URL: https://www.unb.ca/cic/datasets/nsl.html.
4. Sharafaldin I., Lashkari A. H., Ghorbani A. A. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP). 2018. P. 108–116. DOI: https://doi.org/10.5220/0006639801080116.