ВІСНИК ПРИАЗОВСЬКОГО ДЕРЖАВНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ
2025р. Серія: Технічні науки Вип. 52
p-ISSN: 2225-6733; e-ISSN: 2519-271X

УДК 004.75:004.93 DOI: 10.31498/2225-6733.52.2025.350996

РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ГЕНЕРАЦІЇ МЕРЕЖЕВОГО ТРАФІКУ В КОМП'ЮТЕРНИХ МЕРЕЖАХ ДЛЯ ЗАДАЧ КІБЕРБЕЗПЕКИ

Проніна О.І. *канд. техн. наук, доцент, ДВНЗ «Приазовський державний технічний університет», м. Дніпро, ORCID: https://orcid.org/0000-0001-7085-8027, e-mail: pronina_o_i@pstu.edu;*

Рейжевський М.І. *магістр, ДВНЗ «Приазовський державний технічний університет», м. Дніпро, ORCID: https://orcid.org/0009-0006-8212-0111, e-mail: regevsky03@gmail.com*

У статті досліджено проблему формування якісних навчальних вибірок для задач виявлення кіберзагроз у комп'ютерних мережах, що є критично важливою складовою для подальшого ефективного застосування методів машинного навчання. Проаналізовано сучасні підходи до моніторингу мережевого трафіку, класифікації кібератак та використання інтелектуальних систем виявлення вторгнень. Показано, що використання виключно реального мережевого трафіку суттєво ускладнює процес навчання моделей через обмежений доступ до даних, їх фрагментарність і дисбаланс класів. Запропоновано програмне рішення для генерації синтетичного мережевого трафіку, яке дозволяє моделювати як нормальну мережеву активність, так і різні типи кібератак. Розроблено програмне забезпечення надає автоматизований конвеєр створення датасетів із можливістю масштабування, маркування даних та збереження результатів у стандартних форматах для подальшого використання в алгоритмах машинного навчання. Реалізовано підтримку 13 типів мережевих атак, що охоплюють найбільш поширені сучасні загрози. Особливу увагу приділено математичній моделі підготовки даних, яка включає нормалізацію ознак, поділ вибірки на тренувальну, валідаційну та тестову підмножини, використання k-кратної перехресної перевірки та механізм ін'єкції синтетичних аномалій для усунення дисбалансу класів. Такий підхід дозволяє підвищити узагальнюючу здатність моделей і забезпечити об'єктивну оцінку їх ефективності. Експериментальні результати підтверджують коректність роботи розробленого програмного забезпечення та достатній обсяг і варіативність згенерованих даних. Отримані датасети можуть бути використані для навчання, тестування та порівняння моделей машинного навчання в задачах інтелектуального моніторингу трафіку та виявлення кіберзагроз, що визначає практичну цінність запропонованого підходу.

Ключові слова: *кіберзагрози, генерація мережевого трафіку, програмне забезпечення, виявлення аномалій, машинне навчання.*

Постановка проблеми

Сучасний світ дедалі більше переходить у цифровий формат: бізнес, освіта, медицина та навіть особисте спілкування активно здійснюються онлайн, у мережі Інтернет. Як і у фізичному просторі, у віртуальному середовищі також існують злочинці, які намагаються скористатися вразливістю користувачів чи систем. У результаті питання захисту інформації від кібер шахраїв, здатних викрадати цифрові дані, набуває все більшої актуальності та стає одним із ключових викликів нашого часу.

Кібератаки посідають важливе місце у сучасному інформаційному суспільстві. Вони стали одним із ключових викликів для держав, бізнесу та окремих користувачів, оскільки безпека даних напряму залежить від ефективності захисту інформаційних систем [1].

Кожна країна світу стикається з проблемами кіберзлочинності та витрачає значні ресурси на забезпечення кібербезпеки. Хакерські угруповання, використовуючи вразливості програмного забезпечення та соціальну інженерію, здійснюють атаки з метою отримання фінансової вигоди, викрадення даних або порушення роботи критично важливої інфраструктури. Такі атаки стають інструментом як економічного тиску, так і політичної боротьби у глобальному масштабі.

Одним із ключових аспектів вивчення кібербезпеки є аналіз трафіку. Мережевий трафік [2] відображає всі процеси обміну даними між користувачами, серверами та різними пристроями. Саме у трафіку можна виявити шкідливу активність, аномалії або ознаки атак. Трафік ділиться на вхідний та вихідний, і для кожного з них використовуються методи моніторингу, фільтрації та шифрування. Сучасні системи аналізу трафіку (наприклад, IDS/IPS) дозволяють автоматично визначати підозрілу активність і блокувати потенційні загрози.

Існує велика кількість типів кібератак серед найпоширеніших можна виділити [3]:

DDoS-атаки – перевантаження серверів великою кількістю запитів з метою вивести ресурс з ладу;

Фішинг – викрадення особистих даних користувачів через підроблені сайти чи електронні листи;

Вірусні атаки та шкідливе ПЗ – зараження системи програмами, які викрадають інформацію або знищують дані;

SQL-ін'єкції – спроби отримати доступ до баз даних шляхом вставки шкідливих запитів;

Атаки нулевого дня – використання ще невідомих вразливостей у програмному забезпеченні.

Аналіз останніх досліджень та публікацій

Збільшення числа комп'ютерних інцидентів, пов'язаних із зовнішнім втручанням у роботу системи, спонукало дослідників до розробки системи

Copyright © 2025, Authors. This is an open access article under the Creative Commons CC BY license

p-ISSN: 2225-6733; e-ISSN: 2519-271X

своєчасного втручання. Сьогодні такі системи стали необхідним компонентом інфраструктури безпеки організацій, виявлення і передування також є складовою повної роботи фахівців з кібербезпеки. Дана стаття [4] присвячена дослідженню технологій комп'ютерних атак, огляду системи виявлення вторгнень, методів аналізу виявлених атак, системи запобігання вторгнень. Авторами наведено і проаналізовано класифікацію, компоненти і архітектуру системи IDS. Запропоновано підходи до захисту комп'ютерної мережі на базі виявлення системи вторгнення.

Розробка інформаційних технологій виявлення атак є актуальною задачею. Наприклад, у статті [5] розглянуто інформаційну технологію інтелектуального моніторингу трафіку комп'ютерних мереж для виявлення атак та аномальної поведінки. Проаналізовано сучасні підходи до моніторингу мережевого трафіку, методи виявлення аномалій і застосування машинного навчання для аналізу та розвитку алгоритмів виявлення загроз. Здійснено класифікацію систем інтелектуального моніторингу, наведено порівняльну характеристику їх переваг і недоліків та окреслено перспективи підвищення ефективності виявлення мережевих атак.

Віддалений доступ та передача даних роблять мережі вразливими до різних загроз. Таким чином, мережева безпека є важливою для обміну даними та комунікації. У статті [6] наголошується на мережевій безпеці в IT-системах та розглядаються сучасні мережеві загрози та засоби захисту. Головна мета – запобігти доступу хакерів до захищених даних та забезпечити безпечний комунікаційний пристрій для споживачів.

Дослідження [7] присвячено використанню математичних методів, таких як теорія ймовірностей, теорія ігор, графові моделі та статистичні підходи для побудови моделей, що дозволяють відтворювати динаміку загроз у реальних мережах. Методологія базується на моделюванні різних сценаріїв атак, що впливають на інформаційну безпеку. Використання цих моделей дає змогу створювати точні алгоритми для запобігання атакам, що, у свою чергу, забезпечує надійність та безпеку мережевої інфраструктури. Це дослідження надає унікальну можливість глибше зрозуміти природу кібератак, що робить його цінним ресурсом для фахівців з безпеки.

В роботі [8] розроблено метод генерації зразків мережевих атак на основі глибоких згорткових генеративно-змагальних мереж (DCGAN) та метод захисту від змагальних зразків на основі багатомасштабних GAN, а також виконується перевірка практичності цих двох методів за допомогою експериментів. Наведено порівняння трьох методів генерації змагальних зразків: AE-CDA, AE-DEEP та AE-ATTACK, метод генерації змагальних зразків на основі DCGAN у цій статті може ефективніше впливати на функцію виявлення моделі виявлення аномалій, має кращу стабільність та універсальність, а також може підтримувати відносно стабільний ефект атаки на широкому діапазоні моделей та наборів даних.

Робота [9] є цікавим оглядом досліджень мережевої безпеки, класифікації різних атак та загроз, а також заходів, які необхідно впровадити для захисту. У статті також описано різні концепції, пов'язані з безпекою, включаючи мережеву безпеку, криптографію та шифрування.

Робота [10] присвячена питанню бездротових мереж, що стали невід'ємною частиною повсякденного життя, забезпечуючи підключення для широкого кола пристроїв. Однак, зі зростанням використання бездротових мереж, безпека стала серйозною проблемою. Атаки можуть призвести до втрати конфіденційності, цілісності та доступності мережевих ресурсів. Тому в цьому дослідженні досліджувалися вразливості та проблеми, пов'язані з MAC-адресами, підміною DHCP та несанкціонованими атаками SSH. Аналіз рішення для запобігання та пом'якшення цих атак було проведено за допомогою мережевого моделювання з використанням Cisco Packet Tracer Windows версії 8.1.1. Проект системи мав забезпечити зручний інтерфейс для мережевих адміністраторів для моніторингу своїх мереж та перевірки на наявність аномалій.

В роботі [11] проводиться порівняльний аналіз системи виявлення атак та запобігання вторгненням. Наукова новизна дослідження досягнута у визначенні ефективності розроблюваної системи шляхом вирішення завдань багатокритеріального прийняття рішень. У роботі виконується декомпозиція задачі прийняття рішень із виділенням головної цілі та альтернативою з використанням методу Сааті. Результати проведеного аналізу підтверджують, що розроблена система є ефективною та актуальною.

Огляд розширених типів мережевих атак є доцільним дослідженням, що приведено в роботі [12]. Для кожного типу атаки описано джерела та об'єкти атаки, мету та результати атаки, дії, що застосовуються для досягнення мети атаки. Визначено умови можливості здійснення кожного типу атаки. На базі знайдених відомостей зібрано та класифіковано чинники, які дозволяють успішно виконати атаки. Окреслено напрями посилення стійкості до мережевих атак.

У зв'язку з тим, що навчання моделей на реальному трафіку є проблемною задачею, питання створення застосунку для генерації вибірки де є різні типи погроз для мережі є актуальним питанням. Для подальшого навчання моделей машинного навчання необхідно мати якісну вибірку, щоб підвищити якість розпізнавання кібер погроз. Аналіз літератури показав, що створення якісного матеріалу для навчання моделей є перспективним завданням.

Мета статті

Розробка програмного рішення, яке забезпечить генерацію різного типу трафіку, що включає кібератаки та нормальний трафік для створення вибірки

p-ISSN: 2225-6733; e-ISSN: 2519-271X

даних. Отримана вибірка в подальшому буде використовуватися для навчання на моделях машинного навчання для виявлення кіберзагроз.

Виклад основного матеріалу

У поставленій задачі важливим етапом є формування якісних навчальних вибірок, оскільки саме від їх структури залежить ефективність моделей машинного навчання. Дані для задачі виявлення аномалій, класифікації загроз та прогнозування активності можуть походити з різних джерел: журналів подій (logs), мережевих потоків (network flows), системних лічильників продуктивності, а також штучно згенерованих сценаріїв. Для підвищення надійності використовується комбінований підхід – реальні дані доповнюються синтетичними профілями «норми» та «аномалій».

Принцип розбиття загальної вибірки є важливим етапом, що включає в себе поділ вибірки на підмножини. Для навчання та перевірки якості моделей дані розділяються на тренувальну, валідаційну та тестову підмножини. Це дозволяє уникнути перенавчання та оцінити узагальнюючу здатність алгоритму:

$$X_{train}, Y_{train}, X_{test}, Y_{test}, X_{val}, Y_{val} = split(X, y, test_{size} = \alpha), \quad (1)$$

де X – матриця ознак,

Y – вектор класів (норма/аномалія),

α – частка тестової вибірки (зазвичай 0.2–0.3),

y – вектор цільових значень (labels), наприклад, класи «норма» / «аномалія»,

X_{train}, Y_{train} – підмножина даних, що використовується для навчання моделі,

X_{val}, Y_{val} – валідаційна вибірка, яка використовується для налаштування параметрів моделі й запобігання перенавчанню,

X_{test}, Y_{test} – тестова вибірка, застосовується для фінальної перевірки якості після навчання.

Для більш надійної оцінки використовується метод k-кратної перехресної перевірки. Вибірка розбивається на k підмножин і кожна з них послідовно виступає як тестова. Остаточна оцінка моделі визначається як середнє значення по всім ітераціям:

$$CV = \frac{1}{k}\sum_{j=1}^{k} Score_j, \quad (2)$$

де $Score_j$ – значення метрики на j-тій ітерації.

Для покращення якості навчання моделей, виявлення відхилень передбачено етап генерації синтетичних аномалій. Цей етап дозволяє доповнити наявний набір даних штучно створеними прикладами рідкісних подій, що імітують потенційні аномалії у реальному середовищі. Такий підхід забезпечує більш збалансований розподіл класів та підвищує здатність алгоритму розпізнавати нетипові ситуації. Оскільки дані в реальних системах часто мають дисбаланс (аномалії зустрічаються рідко), вводиться механізм ін'єкції синтетичних прикладів. Вони формуються на основі відхилень у багатовимірному просторі ознак, що дозволяє контролювати частку позитивного класу P_{pos} і складність задачі:

$$D_{aug} = D_{real} = D_{synthetick}(P_{pos,\delta}), \quad (3)$$

де D_{real} – реальні дані,

$D_{synthetick}$ – згенеровані точки з відхиленням δ,

P_{pos} – бажана частка позитивного класу.

Оскільки різні ознаки мають різний розмір, то доцільним є використання нормалізації. Нормалізація (або стандартизація) дозволяє привести всі вхідні параметри до єдиної шкали, що особливо важливо при роботі з алгоритмами, чутливими до масштабів ознак. Формула нормалізаціє має наступний вигляд:

$$X_{norm} = \frac{x-\mu}{\sigma}, \quad (4)$$

де x – початкове значення ознаки;

μ – середнє значення ознаки у вибірці,

σ – стандартне відхилення.

Таким чином, побудована математична модель підготовки даних охоплює ключові етапи: стандартизацію ознак, поділ на підмножини, використання кросвалідації та ін'єкцію синтетичних аномалій. Це забезпечує збалансовану навчальну базу та створює умови для об'єктивної оцінки алгоритмів машинного навчання в задачах виявлення загроз у комп'ютерних мережах.

Розроблена математична модель лягла в основу першої частини системи виявлення аномалій, а саме генерації трафіку.

Система моніторингу та виявлення аномалій реалізована у вигляді системи з веб-інтерфейсом [13], що забезпечує зручність доступу, мобільність використання та можливість віддаленого контролю стану комп'ютерної мережі.

Методологія генерації трафіку включає в себе наступні етапи: визначення типу запису, генерація даних, додавання ML-ознак, маркування, збереження у форматі CSV.

Інтерфейс розробленої першої частини системи моніторингу та виявлення аномалій у комп'ютерних мережах представлено на рисунку 1.

Користувач вибирає потрібні параметри для генерування маркованих і немаркованих дата сетів.

За замовчуванням тривалість генерації вибірки складає 30 хвилин, швидкість генерації приблизно 100 пакетів в секунду. Для створення реалістичних даних частота атак складає 5 атак за хвилину. Тобто загальна кількість при генерації за замовченням близько 180 000

При цьому розподіл трафіку має наступний вигляд: нормальний трафік – 70% (≈126 000 записів); атаки – 20% (≈36 000 записів); атакований трафік – 10% (≈18 000 записів).

p-ISSN: 2225-6733; e-ISSN: 2519-271X

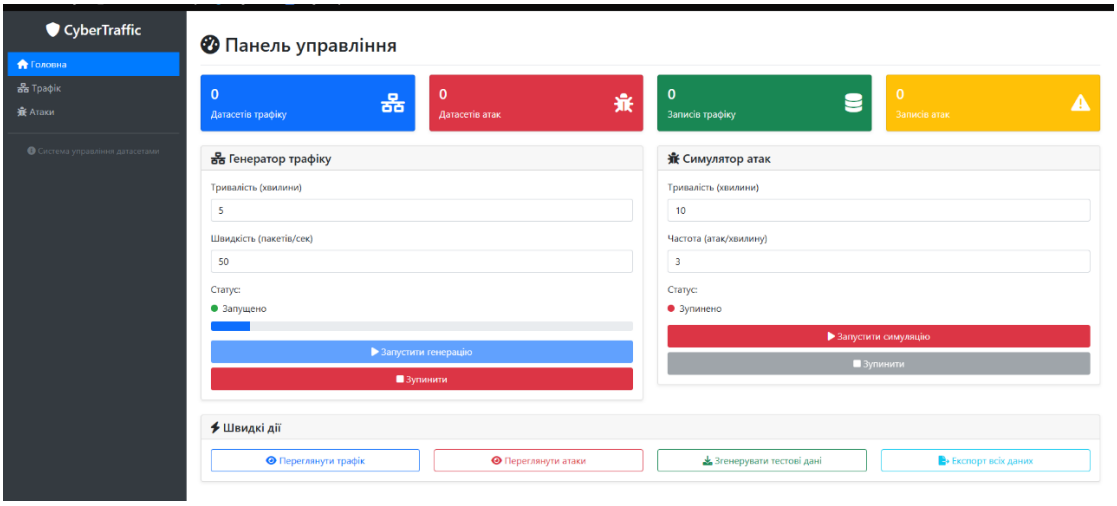

Рис. 1 – Інтерфейс розробленого програмного забезпечення з генерації трафіку

Розроблене програмне забезпечення дозволяє виконувати масштабованість вибірок. Так малі датасети: 1-5 хвилин (6 000-30 000 записів); середні датасети: 10-30 хвилин (60 000-180 000 записів); великі датасети: понад 60 хвилин (360 000+ записів).

Отримана вибірка (датасет) має макровану частину, що використовується для навчання моделей машинного навчання, та немаркований датасет, який застосовується для тестування.

В програмному забезпеченні генерації підтримується генерація 13 типів кібератак, зокрема DDoS, Port Scan, SQL Injection, XSS, Brute Force, Malware, Phishing, Man-in-the-Middle, Buffer Overflow, Privilege Escalation, Data Exfiltration, Ransomware та Unknown [9].

Атакований трафік має свої особливості, що були виявлені при дослідженні результатів генерації, а саме менший розмір пакетів, обмежений набір протоколів, низький TTL, підозрілі TCP прапори.

Оскільки генерація трафіка включає в себе не лише атакований трафік, а ще й нормальний трафік, тому реалізовано характеристики нормального трафіку: протоколи: TCP, UDP, ICMP, HTTP, HTTPS, DNS, FTP, SSH. Розмір пакетів: 64-1500 байт. TTL: 32-255. TCP прапори: SYN, ACK, FIN, RST, PSH, URG.

Для забезпечення якості та варіативності даних, процес створення навчальних вибірок був автоматизований. Розроблений конвеєр (pipeline) генерації атак дозволяє моделювати різні сценарії загроз, такі як DoS, SQL-ін'єкції, XSS та Brute Force, з подальшим маркуванням даних для навчання моделей. Схема потоку генерації вибірки атак та приклади команд для запуску симуляцій наведені на рисунку 2.

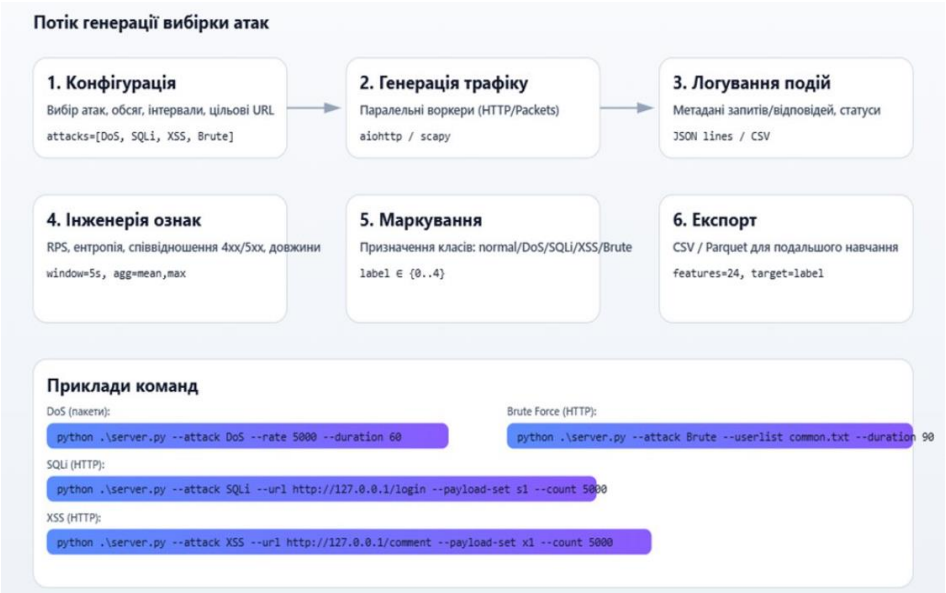

Рис. 2 – Схема потоку генерації синтетичної вибірки атак та приклади скриптів запуску

p-ISSN: 2225-6733; e-ISSN: 2519-271X

Як видно зі схеми на рисунку 2, процес складається з шести етапів: від конфігурації параметрів атаки до експорту готового файлу у форматі CSV або Parquet. Важливим етапом є «Інженерія ознак», де розраховуються статистичні показники (наприклад, ентропія, RPS), які є критичними для коректної роботи алгоритмів машинного навчання. Використання скриптів (server.py) дозволяє гнучко налаштовувати тривалість та інтенсивність атак.

Для перевірки отриманих результатів було виконано експериментальне дослідження, в якому відбувалась генерація трафіку. Кожен раз трафік генерується по різному, що дозволяє робити різні набори вибірок. Приклад розподілу загроз для навчальної та тестової вибріки наведено в таблиці 1. Також в програмному забезпеченні можна зробити розподіл за типами даних та розподіл за наявністю міток та загроз.

Таблиця 1

Розподіл загроз для навчального та тестувального датасетів

Тип загрози	Навчальний датасет	% від навч.	Тест (з мітками)	% від тесту (з мітками)	Тест (без міток)	% від тесту (без міток)	Всього	% від загалу
	4970	100	4975	100	4962	100	14907	100
Нормальний трафік	3,482	70,06	3,441	69,17	2,479	49,96	9,402	63,07
Атакований трафік	474	9,54	564	11,34	0	0	1,038	6,96
Brute_Force	87	1,75	86	1,73	212	4,27	385	2,58
Buffer_Overflow	95	1,91	84	1,69	194	3,91	373	2,50
DDoS	70	1,41	90	1,81	232	4,68	392	2,63
Data_Exfiltration	82	1,65	91	1,83	187	3,77	360	2,41
Malware	95	1,91	73	1,47	212	4,27	380	2,55
Man_in_the_Middle	84	1,69	71	1,43	212	4,27	367	2,46
Phishing	93	1,87	64	1,29	197	3,97	354	2,37
Port_Scan	91	1,83	84	1,69	210	4,23	385	2,58
Privilege_Escalation	78	1,57	87	1,75	189	3,81	354	2,37
Ransomware	81	1,63	84	1,69	224	4,51	389	2,61
SQL_Injection	78	1,57	74	1,49	195	3,93	347	2,33
XSS	80	1,61	82	1,65	219	4,41	381	2,56
	Всього загроз: 1488	29,94	Всього загроз: 1534	30,83	Всього загроз: 2483	50,04	Всього загроз: 5505	36,93

Аналіз структури згенерованих даних дозволяє виявити кілька важливих закономірностей, що підтверджують ефективність обраного підходу. По-перше, у навчальному датасеті дотримано балансу між легітимним трафіком (70,06%) та атаками (29,94%). Це дозволить моделям машинного навчання ефективно вивчати ознаки «норми», маючи при цьому достатню кількість прикладів для кожного з 12 типів загроз. Зокрема, рівномірний розподіл таких складних атак, як SQL Injection (1,57%), Ransomware (1,63%) та Privilege Escalation (1,57%), гарантує, що модель не буде ігнорувати рідкісні події через дисбаланс класів. По-друге, особливу цінність для тестування представляє набір даних без міток, де частка аномалій штучно збільшена до 50,04%.

Такий підхід дозволить перевірити стійкість алгоритмів до інтенсивних сплесків шкідливої активності, що часто спостерігається під час реальних кібератак. По-третє, категорія «Атакований трафік» (9,54% у навчальній вибірці) виступає в ролі сполучного елемента, описуючи стан системи під час перебігу атаки, що є критично важливим для мінімізації помилок другого роду (пропущених атак). Таким чином, розроблений програмний продукт не просто генерує випадкові пакети, а створює структуроване середовище даних, яке відповідає математичній моделі підготовки вибірок (стандартизація, ін'єкція аномалій та кросвалідація). Це забезпечує високу узагальнюючу здатність майбутніх систем виявлення вторгнень та підтверджує практичну цінність запропонованого конвеєра генерації.

Аналізуючи отримані результати можна зробити висновок, що розроблений програмний продукт для генерація трафіку працює коректно. Отримані результати є достатньо об’ємними, щоб можна було в подальшому використовувати їх для навчання моделей машинного навчання визначення кіберзагроз.

Висновки

У ході виконання роботи проаналізовано сучасний стан проблеми кібербезпеки та виявлення мережевих атак, зокрема роль аналізу мережевого трафіку як одного з ключових джерел інформації про шкідливу активність. Проведений огляд наукових публікацій показав, що ефективність інтелектуальних систем виявлення загроз значною мірою залежить від якості та репрезентативності навчальних вибірок.

Розроблено математичну модель підготовки даних для задач машинного навчання, яка охоплює основні етапи формування вибірок: нормалізацію ознак, поділ на підмножини, використання крос-валідації та генерацію синтетичних аномалій. Запропонований підхід дозволяє усунути проблему дисбалансу класів і підвищити здатність моделей розпізнавати рідкісні та нетипові мережеві події.

Створено програмне забезпечення для генерації мережевого трафіку з веб-інтерфейсом, яке підтримує моделювання нормального та атакованого трафіку, масштабування датасетів і генерацію різних типів кібератак. Реалізований конвеєр генерації забезпечує автоматизацію процесу створення навчальних і тестових вибірок, а також формування наборів із мітками та без міток для різних сценаріїв використання.

Експериментальні дослідження підтвердили коректність роботи розробленого програмного продукту та відповідність розподілу згенерованих даних заданим параметрам. Отримані результати свідчать про доцільність використання запропонованої системи для підготовки даних у задачах навчання моделей машинного навчання з виявлення кіберзагроз, а також визначають перспективи подальшого розвитку системи в напрямі інтеграції з інтелектуальними засобами моніторингу комп’ютерних мереж.

Перелік використаних джерел

[1] Steinberg J. Cybersecurity For Dummies. Hoboken, NJ : John Wiley & Sons, Inc., 2019. 368 p.

[2] Bhuyan M. H., Bhattacharyya D. K., Kalita J. K. Network Traffic Anomaly Detection and Prevention: Concepts, Techniques, and Tools. Springer Cham, 2017. 263 p. DOI: https://doi.org/10.1007/978-3-319-65188-0.

[3] Likhith G., Ali J., Vidyashree R. Detection of Cyber Attacks using Artificial Intelligence. International *Journal of Innovative Research in Science, Engineering and Technology*. 2025. Vol. 14, iss. 5. Pp. 13863-13871. DOI: https://doi.org/10.15680/IJIRSET.2025.1405370.

[4] Пашорін В. І., Кравчук П. Ю., Крайчак Є. В. Застосування систем виявлення вторгнень для захисту комп’ютерних мереж. *Збірник наукових праць Європейського університету*. 2023. С. 89-99. DOI: https://doi.org/10.36919/978-966-301-266-7/1.2024.89.

[5] Hasan M., Khan R. Network threats, attacks and security measures: a review. *International Journal of Advanced Computer Research*. 2017. Vol. 9. Pp. 116-120. DOI: https://doi.org/10.26483/ijarcs.v8i9.4641.

[6] Dwivedi A. K., Dwivedi M., Kumar M. Advances in network security: a comprehensive analysis of measures, threats, and future research directions. *International Journal of Emerging Technologies and Innovative Research*. 2023. Vol. 10, iss. 7. Pp. g64-g69.

[7] Doroshenko D. Prediction of Network Threats and Attacks by Mathematical Simulation. *Challenges and Issues of Modern Science*. 2024. Vol. 3. Pp. 173-179.

[8] Shan J., Ma H., Li J. Research on Network Attack Sample Generation and Defence Techniques Based on Generative Adversarial Networks. *Applied Mathematics and Nonlinear Sciences*. 2024. Vol. 9, iss. 1. Pp. 1-20. DOI: https://doi.org/10.2478/amns-2024-3550.

[9] Najem D. F., Kareem S. M. A review on cyber security and cyber attacks. *Journal of Al-Qadisiyah for Computer Science and Mathematics*. 2025. Vol. 17, no. 2. Pp. 153-161. DOI: https://doi.org/10.29304/jqcsm.2025.17.22195.

[10] Security analysis in wireless networks / Rahim M., Jimada-Ojuolape B., Omolara M., Adesina L. *Caliphate Journal of Science and Technology*. 2025. Vol. 7. Pp. 1-11. DOI: https://doi.org/10.4314/cajost.v7i1.01.

[11] Гринченко П. Дослідження розроблюваної системи виявлення мережевих атак (СВМА) із використанням MAI. *Information Technology: Computer Science, Software Engineering and Cyber Security*. 2024. № 2. Pp. 25-33. DOI: https://doi.org/10.32782/IT/2024-2-4.

[12] Годлевський О. Б., Мороховець М. К., Щоголєва Н. М. Аналіз поширених типів мережевих атак та чинники, що уможливлюють їх успішне здійснення. *Information Technologies and Systems*.

2025. No. 2(2). Pp. 55-80. DOI: https://doi.org/10.15407/intechsys.2025.02.055.

[13] Perdana F., Supratman E., Saputra D. Designing a Modern Web Interface with Vue.js and Tailwind for University Information System. *Brilliance: Research of Artificial Intelligence*. 2024. Vol. 4. Pp. 956-963. DOI: https://doi.org/10.47709/brilliance.v4i2.5409.

DEVELOPMENT OF SOFTWARE FOR NETWORK TRAFFIC GENERATION IN COMPUTER NETWORKS FOR CYBERSECURITY TASKS

Pronina O.I. *PhD (Engineering), associate professor, SHEI «Priazovskyi state technical university», Dnipro, ORCID: https://orcid.org/0000-0001-7085-8027, e-mail: pronina_o_i@pstu.edu;*

Reizhevskyi M.I. *M.Sc., SHEI «Priazovskyi state technical university», Dnipro, ORCID: https://orcid.org/0009-0006-8212-0111, e-mail: regevsky03@gmail.com*

The article investigates the problem of forming high-quality training samples for cyber threat detection tasks in computer networks, which is a critically important component in the further effective application of machine learning methods. Modern approaches to monitoring network traffic, classifying cyber attacks, and using intelligent intrusion detection systems are analyzed. It is shown that using only real network traffic significantly complicates the process of training models due to limited access to data, its fragmentation, and class imbalance. A software solution is proposed for generating synthetic network traffic, which allows modeling both normal network activity and various types of cyber attacks. The software developed provides an automated pipeline for creating datasets with the ability to scale, label data, and save results in standard formats for further use in machine learning algorithms. Support for 13 types of network attacks, covering the most common modern threats, is implemented. Particular attention is paid to the mathematical model of data preparation, which includes feature normalization, sample division into training, validation and test subsets, the use of k-fold cross-validation and the mechanism of injection of synthetic anomalies to eliminate class imbalance. This approach allows to increase the generalization ability of models and provide an objective assessment of their effectiveness. Experimental results confirm the correctness of the work of the developed software and the sufficient volume and variability of the generated data. The obtained datasets can be used for training, testing and comparison of machine learning models in the tasks of intelligent traffic monitoring and cyber threat detection, which determines the practical value of the proposed approach.

***Keywords**: cyber threats, network traffic generation, software, anomaly detection, machine learning.*

References

[1] J. Steinberg, *Cybersecurity For Dummies*. Hoboken, NJ : John Wiley & Sons, Inc., 2019.

[2] M.H. Bhuyan, D.K. Bhattacharyya, and J.K. Kalita, *Network Traffic Anomaly Detection and Prevention: Concepts, Techniques, and Tools*. Springer Cham Publ., 2017. **doi: 10.1007/978-3-319-65188-0**.

[3] G. Likhith, J. Ali, and R. Vidyashree, "Detection of Cyber Attacks using Artificial Intelligence," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 14, iss. 5, pp. 13863-13871, 2025. **doi: 10.15680/IJIRSET.2025.1405370**.

[4] V.I. Pashorin, P.Yu. Kravchuk, and Ye.V. Kraichak, "Zastosuvannia system vyiavlennia vtorhnen dlia zakhystu kompiuternykh merezh" ["Application of intrusion detection systems to protect computer networks"], *Zbirnyk naukovykh prats Yevropeiskoho universytetu – Collection of scientific papers of the European University*, pp. 89-99, 2023. **doi: 10.36919/978-966-301-266-7/1.2024.89**.

[5] M. Hasan, and R. Khan, "Network threats, attacks and security measures: a review," *International Journal of Advanced Computer Research*, vol. 9, pp. 116-120, 2017. **doi: 10.26483/ijarcs.v8i9.4641**.

[6] A.K. Dwivedi, M. Dwivedi, and M. Kumar, "Advances in network security: a comprehensive analysis of measures, threats, and future research directions," *International Journal of Emerging Technologies and Innovative Research*, vol. 10, iss. 7, pp. g64-g69, 2023.

[7] D. Doroshenko, "Prediction of Network Threats and Attacks by Mathematical Simulation," *Challenges and Issues of Modern Science*, vol. 3, pp. 173-179, 2024.

[8] J. Shan, H. Ma, and J. Li, "Research on Network Attack Sample Generation and Defence Techniques Based on Generative Adversarial Networks," *Applied Mathematics and Nonlinear Sciences*, vol. 9, iss. 1, pp. 1-20, 2024. **doi: 10.2478/amns-2024-3550**.

[9] D.F. Najem, and S.M. Kareem, "A review on cyber security and cyber attacks," *Journal of Al-Qadisiyah for Computer Science and Mathematics*, vol. 17, no. 2, pp. 153-161, 2025. **doi: 10.29304/jqcsm.2025.17.22195**.

[10] M. Rahim, B. Jimada-Ojuolape, M. Omolara, and L. Adesina, "Security analysis in wireless networks," *Caliphate Journal of Science and Technology*, vol. 7, pp. 1-11, 2025. **doi: 10.4314/cajost.v7i1.01**.

[11] P. Hrynchenko, "Doslidzhennia rozrobliuvanoi systemy vyiavlennia merezhevykh atak (SVMA) iz vykorystanniam MAI" ["Research of the network attacks detection system (NADS) under development using MAI"], *Information Technology: Computer Science, Software Engineering and Cyber Security*, № 2, pp. 25-33, 2024. **doi: 10.32782/IT/2024-2-4**.

[12] A.B. Godlevsky, M.K. Morokhovets, and N.M. Shchogoleva, "Analiz poshyrenykh typiv merezhevykh atak ta chynnyky, shcho umozhlyvliuiut yikh uspishne zdiisnennia" ["Analysis of Common Types of Network Attacks, and Factors Enabling their Successful Implementation"], *Information Technologies and Systems*, no. 2(2), pp. 55-80, 2025. **doi: 10.15407/intechsys.2025.02.055**.

[13] F. Perdana, E. Supratman, and D. Saputra, "Designing a Modern Web Interface with Vue.js and Tailwind for University Information System," *Brilliance: Research of Artificial Intelligence*, vol. 4, pp. 956-963, 2024. **doi: 10.47709/brilliance.v4i2.5409**.

Стаття надійшла 10.10.2025
Стаття прийнята 02.11.2025
Стаття опублікована 29.12.2025

Цитуйте цю статтю як: Проніна О. І., Рейжевський М. І. Розробка програмного забезпечення генерації мережевого трафіку в комп'ютерних мережах для задач кібербезпеки. *Вісник Приазовського державного технічного університету. Серія: Технічні науки*. 2025. Вип. 52. С. 75-82. DOI: https://doi.org/10.31498/2225-6733.52.2025.350996.