DOI 10.31891/2307-5732-2024-333-2-15
УДК 004.383.8:681.2

БОНДАРЕНКО АНТОН
Київський національний університет технологій та дизайну
https://orcid.org/0009-0007-5087-6173
e-mail: anton.bondarenko.ua@gmail.com
СТАЦЕНКО ВОЛОДИМИР
Київський національний університет технологій та дизайну
https://orcid.org/0000-0002-3932-792X
e-mail: statsenko.v@knutd.edu.ua

ВИКОРИСТАННЯ МЕТОДІВ ТА МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ПОКРАЩЕННЯ ЕКСПЕРТНИХ СИСТЕМ ВИЯВЛЕННЯ ВТОРГНЕНЬ

У сфері кібербезпеки ефективність систем виявлення вторгнень (IDS) має вирішальне значення для проактивної ідентифікації та пом'якшення кіберзагроз. Це дослідження окреслює нову парадигму для підвищення точності IDS шляхом інтеграції передових методологій штучного інтелекту (AI), тим самим встановлюючи новий стандарт у механізмах захисту мережевої безпеки. Проаналізовано і розроблено систему виявлення вторгнень з використанням новітніх методів машинного навчання.

Ключові слова: варіаційний кодувальник, розпізнавання аномалій, екстремальне градієнтне підсилення, класифікація, експертні системи, системи виявлення вторгнень.

BONDARENKO ANTON M., STATSENKO VOLODYMYR V.
Kyiv National University of Technologies and Design

USE OF ARTIFICIAL INTELLIGENCE METHODS AND MODELS FOR IMPROVING EXPERT SYSTEMS OF INTRUSION DETECTION

In the domain of cyber security, the efficacy of Intrusion Detection Systems (IDS) is critical for the proactive identification and mitigation of cyber threats. This research delineates a novel paradigm for enhancing IDS accuracy through the integration of advanced Artificial Intelligence (AI) methodologies, thereby setting a new benchmark in network security defense mechanisms. Utilizing a synergistic approach that combines both descriptive and inferential statistical analyses, this study introduces an expert system endowed with the capability to detect network intrusions with an unparalleled accuracy rate of 99.98%. By incorporating Extreme Gradient Boosting (XGBoost) for the classification of predefined attack vectors and a Variational Autoencoder (VAE) for anomaly detection, the research extends the boundaries of current cyber threat detection frameworks. These methodologies not only enhance the precision of threat categorization but also introduce a mechanism for the system to adapt to novel, previously unidentified cyber threats through real-time learning and adaptation to emerging data patterns. Critically, the expert system is engineered to facilitate high-speed data processing and supports online learning, making it optimally suited for application in high-traffic network environments. The scientific novelty of this research is encapsulated in the formulation of advanced AI-driven models for the dual purposes of traffic anomaly detection and the classification of cyber-attack types based on distinctive behavioral characteristics. These models are meticulously designed to evolve, learning from new data in real time, thereby continuously enhancing the system's efficacy. In practical terms, the system provides a robust solution for the protection of digital ecosystems against intrusions, enabling the automatic filtration of malicious network traffic. Beyond its immediate applicability, the study contributes to the field of cyber security by laying down a foundational framework for the future exploration of AI-based security solutions. It invites further scientific inquiry into the development of adaptive, intelligent IDS mechanisms, potentially revolutionizing the approach to cyber defense strategies.

Keywords: variational autoencoder, anomaly detection, extreme gradient boosting, classification, expert system, intrusion detection system.

Постановка проблеми

Останніми роками інформаційні ресурси стають дедалі вразливішими до різноманітних кіберзагроз, причому зростає [1] частота та складність атак розподіленої відмови в обслуговуванні (DDoS) і спроб вторгнення. Ці інциденти не тільки порушують нормальне функціонування онлайн-сервісів, але й створюють значні ризики для безпеки конфіденційних даних і критичної інфраструктури. Необхідність захисту цифрових систем від таких загроз ніколи не була настільки критичною, що підкреслює важливість дослідження та вдосконалення систем виявлення вторгнень (IDS) [2]. IDS відіграють ключову роль у виявленні та запобіганню несанкціонованого доступу або атак, тим самим підвищуючи стійкість мереж до кіберзагроз.

Інтеграція алгоритмів машинного навчання, таких як градієнтне підсилення (XGBoost) [3] та варіаційний автокодувальник (VAE) [4], у фреймворки IDS, є значним прогресом у виявленні та запобіганні кібератак. XGBoost із високою ефективністю, масштабованістю та здатністю обробляти розріджені дані вправно класифікує та прогнозує потенційні загрози на основі історичних даних. З іншого боку, VAE за допомогою своїх генеративних можливостей можуть моделювати складні розподіли моделей атак, полегшуючи виявлення нових загроз, які звичайні методи можуть пропустити.

Незважаючи на суттєвий прогрес, якого було досягнуто в останні роки, ця сфера залишається повною проблем, зокрема у виявленні складних незначних вторгнень та адаптації до динамічної природи кіберзагроз. Недавні дослідження почали вирішувати ці проблеми, використовуючи сильні сторони XGBoost і VAE для підвищення точності виявлення та зменшення помилкових спрацьовувань. Однак залишається значний пробіл у дослідженнях, особливо в інтеграції цих технологій у цілісну IDS у реальному часі, яка може ефективно адаптуватися до нових загроз. Метою дослідження є визначення ефекту поєднання XGBoost і VAE в новій структурі IDS на швидкість та точність виявлення атак.

Аналіз останніх досліджень

Розвиток систем виявлення вторгнень значною мірою залежить від використання даних і моделей. Сучасні дослідження в цій галузі акцентують на важливості розробки якісних наборів даних та ефективності моделей машинного навчання.

Створення та характеризація нових наборів даних для IDS, як це зазначено у дослідженнях [5, 6], є критично важливими для підвищення ефективності детекції. Ці дослідження виявили проблеми з існуючими наборами даних та запропонували методи та інструменти для їх збору [7, 8] та вдосконалення, підкреслюючи необхідність якісних даних для тренування передових IDS.

Дослідження, присвячені ефективності моделей в IDS, такі як використання AdaBoost та XGBoost [9–12], а також дослідження [13–16], які зосереджені на застосуванні глибокого навчання в IDS, вказують на значний потенціал інтеграції складних алгоритмів для підвищення точності та ефективності систем виявлення вторгнень. Використання VAE для детекції аномалій і XGBoost для класифікації атак стає основою для створення IDS, що обіцяє значне покращення в роботі з сучасними кіберзагрозами.

Постановка задачі

Проаналізувати інтеграцію eXtreme Gradient Boosting (XGBoost) і Variation Autoencoders (VAE) як систему виявлення вторгнень (IDS), спрямовану на посилення виявлення кіберзагроз у реальному часі та їх запобіганню. Зокрема, це дослідження спрямоване на використання XGBoost для класифікації та ідентифікації відомих видів атак, одночасно використовуючи VAE для виявлення нових, раніше невідомих загроз, автоматизуючи навчання. У дослідженні запропоновано механізм онлайн-навчання, який сприяє динамічному та постійному оновленню моделі IDS. Цей механізм потрібен для того, щоб система могла адаптуватися в режимі реального часу до кіберзагроз, що постійно змінюються, без необхідності повного перенавчання моделі. Отже, цей підхід скорочує час простою системи та підвищує точність виявлення загроз.

Результати дослідження

Дослідження базується на датасеті CIC IDS 2017 [5] та його модифікації [6], що включає правки до роботи програми CICFlowMeter [7, 8] та більш гранулярне маркування даних. Цей датасет часто використовується в дослідженнях роботи систем виявлення вторгнення [9, 10, 13, 14]. Він включає в себе мережеві пакети та потоки. Потоки є статистичними даними кожного з'єднання, зібрані завдяки інструменту CICFlowMeter [10, 11]. В цьому дослідженні для аналізу були обрані потоки через те, що вони в собі містять інформацію за весь період з'єднання, що дає більшу статистичну значимість. З більш детальним описом ознак та промаркованих класів трафіку можна ознайомитись за відповідними ресурсами [5, 6].

Перед початком роботи з датасетом, було побудовано теплокарту кореляцій ознак, які можна побачити на рис. 1. Її було побудовано для аналізу ознак та їх кореляцій щодо промаркованого класу трафіку. Також це було зроблено для валідації якості доданих ознак, які наведені нижче.

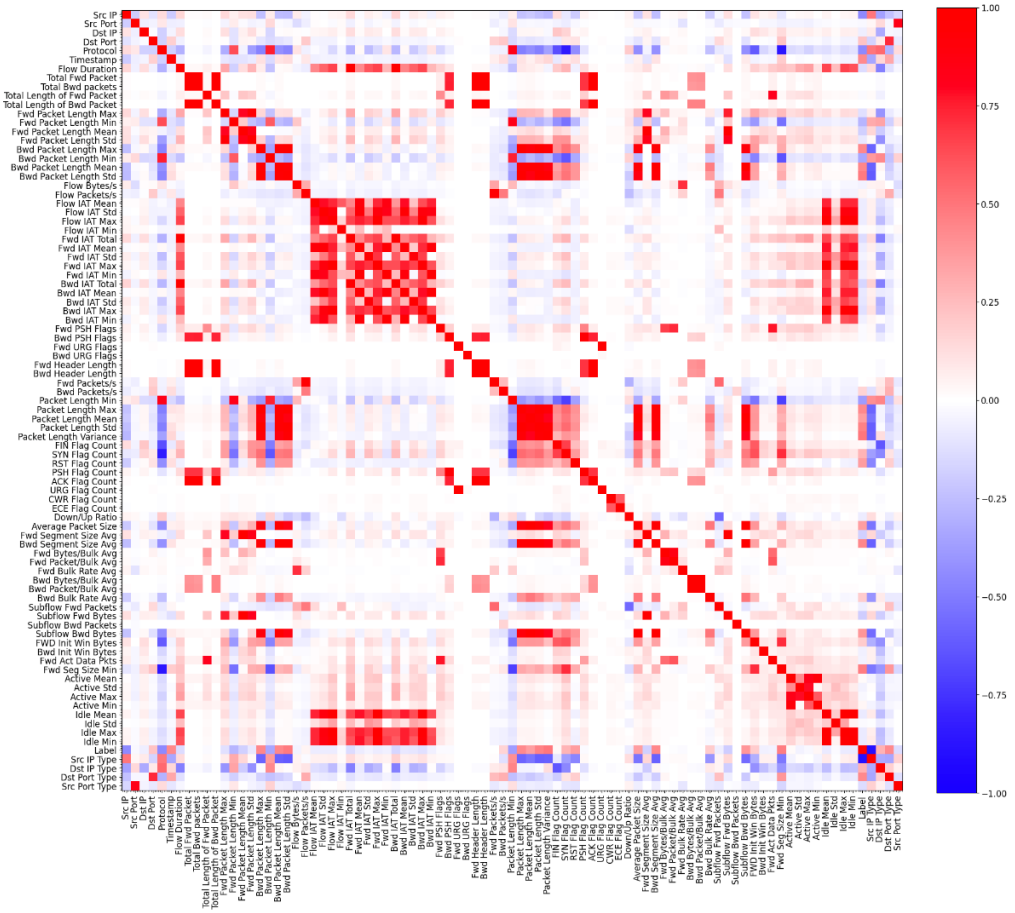


Рис. 1. Теплокарта кореляцій ознак датасету

Під час підготовки датасету для навчання моделі було прибрано наступні ознаки:

- Flow ID – ідентифікатор з'єднання. Це внутрішній ідентифікатор програми CICFlowMeter, який не має ніякого відношення до трафіку.
- Timestamp – дата підключення. Цей параметр було прибрано, щоб моделі не мали упереджень щодо часу відправленого трафіку.
- Src IP – IP адреса відправника. Цей параметр було прибрано, щоб модель не мала упередженості щодо конкретних IP адрес.
- Src Port – IP порт відправника. Цей параметр було прибрано, щоб модель не мала упередженості щодо конкретних IP портів.
- Dst Port – IP порт отримувача. Цей параметр було прибрано, щоб модель не мала упередженості щодо конкретних IP портів.
- Dst IP – IP адреса отримувача. Цей параметр було прибрано, щоб модель не мала упередженості щодо конкретних IP адрес.

Також до датасету були додані наступні ознаки:

- Dst Port Type – тип IP порту відправника
- Src Port Type – тип IP порту отримувача
- Src IP Type – тип IP відправника
- Dst IP Type – тип IP отримувача

Тип IP адреси – значення, що вказує чи є IP адреса локальною до мережі на якій відстежується трафік. Тип IP порту – флаг, що вказує на вид порту за ознакою цільового призначення [17], де порти від 0 до 1023 – системні, від 1024 до 49151 – користувацькі, від 49152 до 65535 – динамічні і як видно на теплокарті кореляцій мають більшу статистичну значимість ніж колонки з яких вони були сформовані, також мають більший взаємозв'язок з промаркованим класом трафіку, що можна побачити в таблиці 1.

Таблиця 1

Порівняння коефіцієнтів кореляції для запропонованих ознак

Назва ознаки	Коефіцієнт кореляції з класом трафіку
Src IP	-0,44
Src Port	-0,02
Dst IP	-0,06
Dst Port	0,22
Src IP Type	-0,90
Dst IP Type	0,28
Dst Port Type	0,43
Src Port Type	-0,13

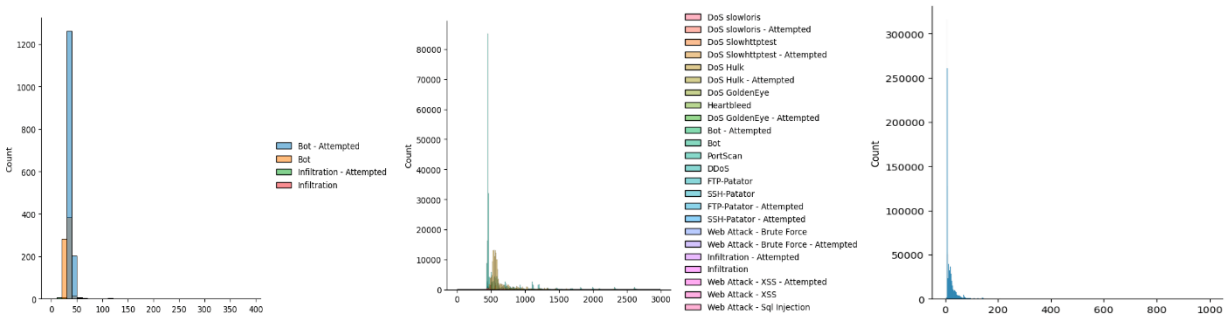
Така різниця коефіцієнтів для IP адрес зумовлена тим що, недоброякісний трафік частіше приходить з зовнішніх мереж, також це дозволить моделям не мати упередженості щодо IP адрес. Щодо типу IP порту, то вірогідність недоброякісного трафіку з системних та користувацьких портів значно більша, ніж з динамічних, що зумовлено особливостями стандартних налаштувань операційних систем [17, 18, 19].

Наступним кроком була розробка моделі виявлення аномалій. Як базова модель був обраний саме VAE, тому що має низку якостей, які його виділяють перед звичайними автокодувальниками (AE) [20] особливо для задач виявлення аномалій [15, 16]. Ця модель з її імовірнісним підходом до вивчення латентних уявлень за своєю суттю є більш гнучкою в моделюванні складних розподілів даних. Ця імовірнісна структура дозволяє VAE краще моделювати невизначеність і мінливість мережевого трафіку, що має вирішальне значення для точного виявлення аномалій, які відхиляються від нормальних моделей.

Модель була навчена виключно на даних, позначених як доброякісні (BENIGN). Така стратегія навчання дозволяє моделі ідентифікувати відхилення від доброякісного трафіку як аномалії. Щоб підвищити надійність моделі на етапі навчання, було застосовано функцію втрат Хубера [21], що робить модель менш чутливою до відхилень у даних та стабілізує навчання, що особливо важливо враховуючи різноманітну природу даних мережевого трафіку. На фазі реконструкції було використано середньо квадратичну похибку (MSE) через її більшу чутливість до різниці в даних.

Також перед тренуванням моделі обов'язково провести нормалізацію даних, що також прискорює та спрощує навчання та задачу пошуку аномалій.

Після навчання варіаційного автокодувальника треба визначити поріг похибки реконструкції за яким буде виявлятись аномалія. Для цього було побудовано гістограми, що зображені на рис. 2 з підрахунком кількості входжень потоків в діапазонах похибки реконструкції, та визначено оптимальне значення порогу – 350.



а б в

Рис 2. Гістограми розподілу похибки реконструкції:
а –розподілу похибки реконструкції Bot та Infiltration атак, б –розподілу похибки реконструкції усіх видів атак, в – розподілу похибки реконструкції доброякісного трафіку

Отже фінальні параметри моделі вийшли наступними:

- Функція похибки на етапі на навчання – функція Хубера
- Функція похибки на етапі використання – середньо квадратична
- Кількість навчальних ітерацій – 20000
- Глибина моделі – 2
- Розмір латентного простору – 5
- Швидкість навчання – 0.01
- Розмір навчального батчу - 256
- Функція активації – ReLU
- Час навчання – 5хв

В результаті навчання було отримано доволі велику точність, яку можна розбити на два види, точність виявлення доброякісного трафіку та точність виявлення аномалій, які мають значення – 99,02% та 91% відповідно, загальна точність - 95,01%.. Більш гранулярну точність по кожному з класів атак представлено в таблиці 2.

Таблиця 2.

Точність роботи варіаційного кодувальника на всьому датасеті

Клас трафіку	**Кількість передбаченого доброякісного трафіку**	**Кількість передбаченого недоброякісного трафіку**	**Точність**
BENIGN	1637718	16209	99.02%
DoS slowloris	0	4001	100%
DoS slowloris - Attempted	0	1706	100%
DoS Slowhttptest	0	1742	100%
DoS Slowhttptest - Attempted	0	3367	100%
DoS Hulk	0	158469	100%
DoS Hulk - Attempted	0	579	100%
DoS GoldenEye	0	7567	100%
Heartbleed	0	11	100%
DoS GoldenEye - Attempted	0	80	100%
Bot - Attempted	1469	0	0%
Bot	694	44	5.96%
PortScan	0	159023	100%
DDoS	0	95123	100%
FTP-Patator	0	3973	100%
SSH-Patator	0	2980	100%
FTP-Patator - Attempted	0	11	100%
SSH-Patator - Attempted	0	8	100%
Web Attack - Brute Force	0	151	100%
Web Attack - Brute Force - Attempted	0	1214	100%
Infiltration - Attempted	16	0	0%
Infiltration	20	12	37.50%
Web Attack - XSS - Attempted	0	652	100%
Web Attack - XSS	0	27	100%
Web Attack - Sql Injection	0	12	100%

Як видно за результатами, модель гарно класифікує аномалії, хоча з видами атак, як Bot або Infiltration впоратись не змогла, бо вони дуже схожі до доброякісного трафіку, що можна побачити в гістограмах розподілу похибки реконструкції рис. 2, також велика кількість доброякісного трафіку була виявлена як аномальна, що може призвести до блокування доступу для нормальних користувачів. Саме тому ця модель не може бути використана, як єдине рішення для системи виявлення вторгнень.

Для системи виявлення вторгнень важливо розуміти не тільки, що трафік доброякісний або ні, але треба знати тип атаки, щоб оператор або автоматична система могли коректно підібрати методи реагування. Саме тому має сенс реалізувати модель класифікації типу атаки. Для цього було використано модель градієнтного підсилення, цей метод класифікації є достатньо відомим та часто використовується в інших дослідженнях систем вторгнення [10, 11, 12].

Отже, цю модель було обрано та навчено на тренувальних даних, і точність була перевірена на всьому датасеті. Через незбалансованість лейблів в датасеті під час тренування використовувались збалансовані ваги. Результат оформлено у вигляді матриці невідповідностей, що можна побачити на рис 3.

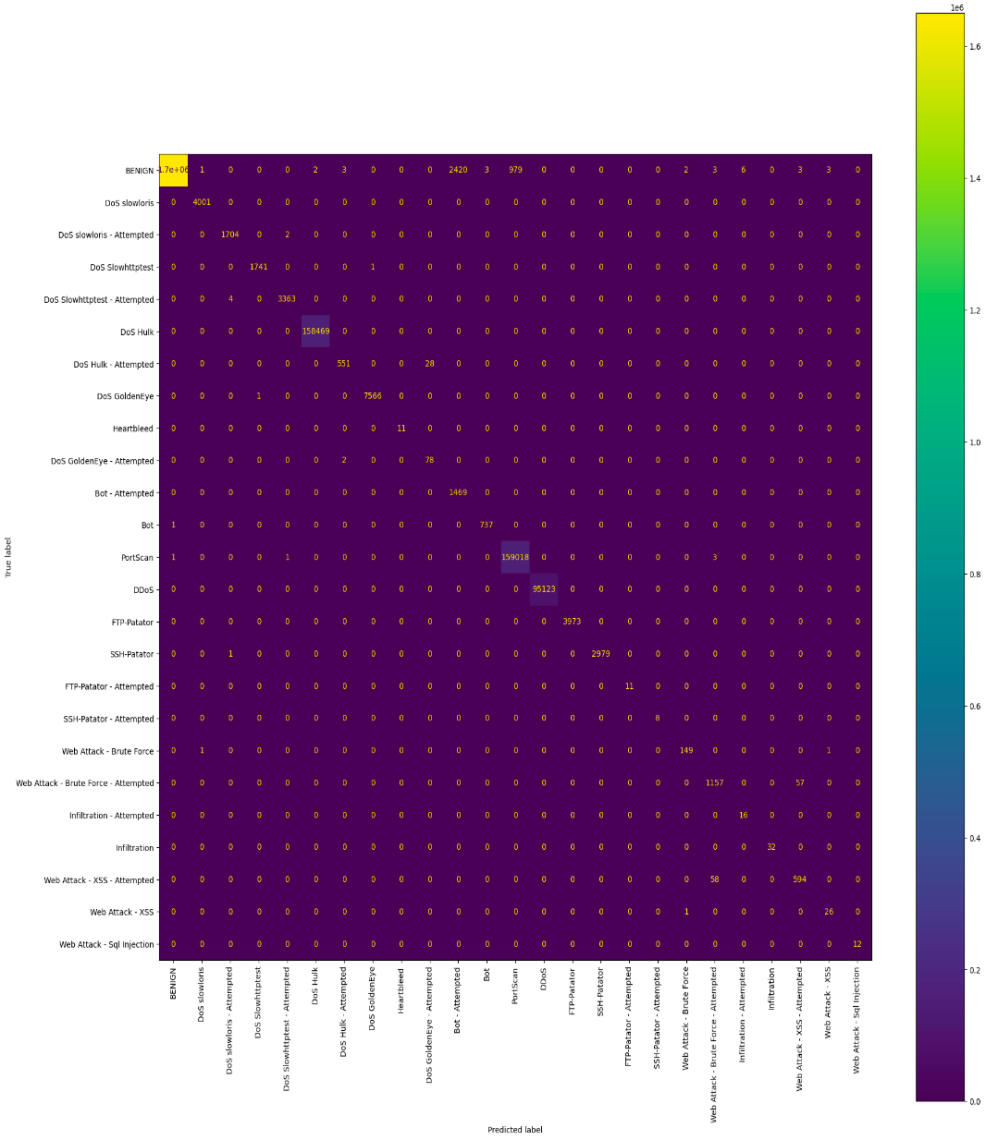


Рис 3. Матриця невідповідностей класифікатора

Згідно з матрицею невідповідностей можна побачити, що ця модель має високу загальну точність – 98,82%, і дуже гарно справляється з усіма видами атак.

Отримавши дві моделі, було побудовано компонент, який поєднує результати цих моделей в єдину систему. Під час аналізу результатів, було виявлено, що обидві моделі гарно доповнюють одна одну і методи атак, які одна модель не виявила, були виявлені іншою. Також було помічено, що детектор аномалій гірше справляється з визначенням доброякісного трафіку, через це було запропоновано наступну таблицю прийняття рішення для трафіку зображену на рис. 4.

Суть такого підходу полягає в тому щоб підвищити точність за допомогою додавання нового типу для визначення трафіку, а саме «підозрілий». Це дозволяє системі більш гранулярно давати відповіді, також користувачам такої системи можна буде використовувати менш жорсткі методи запобігання атакам. Алгоритм роботи та використання такої моделі зазначено на рис. 5.

Рис 4. Матриця прийняття рішення

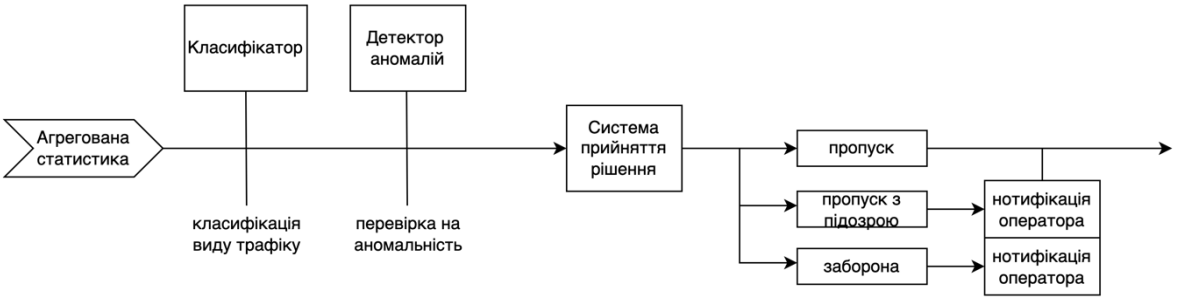


Рис 5. Алгоритм роботи системи ідентифікації вторгнення

Рішення з новим типом для визначення виду трафіку дозволяє досягти більшої точності виявлення недоброякісного трафіку. Результати точності викладені в таблиці 3. Загальна точність складає – 99,98%.

Таблиця 3.

Точність роботи системи на всьому датасеті

Клас трафіку	Кількість передбаченого доброякісного трафіку	Кількість передбаченого недоброякісного трафіку	Кількість передбаченого підозрілого трафіку	Точність
BENIGN	1650502	991	2434	99,79%
DoS slowloris	0	4001	0	100%
DoS slowloris - Attempted	0	1706	0	100%
DoS Slowhttptest	0	1742	0	100%
DoS Slowhttptest - Attempted	0	3367	0	100%
DoS Hulk	0	158469	0	100%
DoS Hulk - Attempted	0	579	0	100%
DoS GoldenEye	0	7567	0	100%
Heartbleed	0	11	0	100%
DoS GoldenEye - Attempted	0	80	0	100%
Bot - Attempted	0	0	1469	100%
Bot	1	44	693	100%
PortScan	0	159022	1	99,86%
DDoS	0	95123	0	100%
FTP-Patator	0	3973	0	100%
SSH-Patator	0	2980	0	100%
FTP-Patator - Attempted	0	11	0	100%
SSH-Patator - Attempted	0	8	0	100%
Web Attack - Brute Force	0	151	0	100%
Web Attack - Brute Force - Attempted	0	1214	0	100%
Infiltration - Attempted	0	0	16	100%
Infiltration	0	11	21	100%
Web Attack - XSS - Attempted	0	652	0	100%
Web Attack - XSS	0	25	2	100%
Web Attack - Sql Injection	0	12	0	100%

Окремо варто приділити увагу онлайн навчанню такої системи. Такий механізм дуже важливий для успішної роботи в реальній системі. Але треба зазначити, що в реалізації адаптивного навчання є невирішені проблеми. VAE, як детектор аномалій дуже легко адаптується до змін, бо він навчається тільки на доброякісному трафіку, тоді як XGBoost потребує більшої уваги, через вимогу в маркуванні даних під кожен вид атаки. У разі додавання нових методів атаки, що дуже вірогідно враховуючи швидкість розвитку кіберзагроз, виникає потреба у створенні додаткових моделей, що будуть класифікувати нові види атак, також потрібна модель, яка буде працювати як роутер, обираючи потрібні моделі для класифікації.

Висновки

1) В результаті дослідження було виявлено, що варіаційний автокодувальник швидко навчається – приблизно 5хв, та дає доволі високу точність – 95.01%, будучи при цьому моделлю навчання без нагляду. Що робить його дуже гарним кандидатом для онлайн навчання та як модель для виявлення аномалій.

2) Було запропоновано нові ознаки в датасет CIC IDS 2017 – Src IP Type, Dst IP Type, Src Port Type, Dst Port Type, які мають наступні коефіцієнти кореляції з промаркованим типом трафіку відповідно - -0.90, 0.28, -0.13, 0.43.

3) Запропоновано матрицю прийняття рішення для виявлення не тільки доброякісного або недоброякісного трафіку, але ще й підозрілого, що дозволяє суттєво підвищити точність виявлення кіберзагроз та дає можливість користувачам системи виконувати менш жорсткі методи уникнення загроз.

4) Створено систему виявлення вторгнень, яка використовує моделі VAE та XGBoost, та надає приголомшливу точність – 99,98%, що покращує точність на 0.02% [21].

Література

1. DDoS threat report for 2023 Q4. *The Cloudflare Blog*, Jan. 09, 2024. URL: https://blog.cloudflare.com/ddos-threat-report-2023-q4

2. Bace R., Mell P. NIST Special Publication on Intrusion Detection Systems Intrusion Detection Systems. 2001. URL: http://cs.uccs.edu/~cchow/pub/ids/NISTsp800-31.pdf

3. Chen T., Guestrin C. XGBoost: a Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16, pp. 785–794, 2016, doi: https://doi.org/10.1145/2939672.2939785.

4. Kingma D. P., Welling M. An Introduction to Variational Autoencoders. Foundations and Trends® in Machine Learning, vol. 12, no. 4, pp. 307–392, 2019, doi: https://doi.org/10.1561/2200000056.

5. Sharafaldin I., Habibi Lashkari A., 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, 2018, doi: https://doi.org/10.5220/0006639801080116.

6. Engelen G., Rimmer V., Joosen W. Troubleshooting an Intrusion Detection Dataset: the CICIDS2017 Case Study. IEEE Xplore, May 01, 2021. URL: https://ieeexplore.ieee.org/abstract/document/9474286

7. Habibi Lashkari A., Draper Gil G., Mamun M. S. I., Ghorbani A. A. Characterization of Tor Traffic using Time based Features. Proceedings of the 3rd International Conference on Information Systems Security and Privacy, 2017, doi: https://doi.org/10.5220/0006105602530262.

8. Applications | Research | Canadian Institute for Cybersecurity | UNB. *www.unb.ca*. URL: https://www.unb.ca/cic/research/applications.html#CICFlowMeter

9. Yulianto A., Sukarno P., Suwastika N. A. Improving AdaBoost-based Intrusion Detection System (IDS) Performance on CIC IDS 2017 Dataset. Journal of Physics: Conference Series, vol. 1192, p. 012018, Mar. 2019, doi: https://doi.org/10.1088/1742-6596/1192/1/012018.

10. D'hooge L., Wauters T., Volckaert B., De Turck F. Classification Hardness for Supervised Learners on 20 Years of Intrusion Detection Data. IEEE Access, vol. 7, pp. 167455–167469, 2019, doi: https://doi.org/10.1109/access.2019.2953451.

11. Dhaliwal S., Nahid A.-A., Abbas R. Effective Intrusion Detection System Using XGBoost. Information, vol. 9, no. 7, p. 149, Jun. 2018, doi: https://doi.org/10.3390/info9070149.

12. Bhati B. S., Chugh G., Al-Turjman F., Bhati N. S. An improved ensemble based intrusion detection technique using XGBoost. Transactions on Emerging Telecommunications Technologies, Aug. 2020, doi: https://doi.org/10.1002/ett.4076.

13. Almaghthawi Yousef, Ahmad I., Alsaadi F. E. Performance Analysis of Feature Subset Selection Techniques for Intrusion Detection. *Mathematics*, vol. 10, no. 24, pp. 4745–4745, Dec. 2022, doi: https://doi.org/10.3390/math10244745.

14. Jose J., Jose D. V. Deep learning algorithms for intrusion detection systems in internet of things using CIC-IDS 2017 dataset. | International Journal of Electrical & Computer Engineering (2088-8708) | EBSCOhost. *openurl.ebsco.com*, Feb. 01, 2023. URL: https://openurl.ebsco.com/contentitem/gcd:160447150?sid=ebsco:plink:scholar&id=ebsco:gcd:160447150&crl=c (accessed Mar. 09, 2023).

15. Liu C., Antypenko R., Sushko I., Zakharchenko O. Intrusion Detection System After Data Augmentation Schemes Based on the VAE and CVAE. IEEE Transactions on Reliability, vol. 71, no. 2, pp. 1000–1010, Jun. 2022, doi: https://doi.org/10.1109/TR.2022.3164877.

16. Esmaeilzadeh S., Salajegheh N., Ziai A., Boote J. Abuse and Fraud Detection in Streaming Services Using Heuristic-Aware Machine Learning. arXiv:2203.02124 [cs], Mar. 2022, URL: https://arxiv.org/abs/2203.02124

17. Cotton M., Eggert L., Touch J. D., Westerlund M., Cheshire S. Internet Assigned Numbers Authority (IANA) Procedures for the Management of the Service Name and Transport Protocol Port Number Registry. IETF, Aug. 01, 2011. URL: https://datatracker.ietf.org/doc/html/rfc6335

18. "IP Sysctl — The Linux Kernel documentation," *www.kernel.org*. URL: https://www.kernel.org/doc/html/latest//networking/ip-sysctl.html#ip-variables (accessed Mar. 09, 2024).

19. Deland-Han. The default dynamic port range for TCP/IP has changed in Windows Vista and in Windows Server 2008 - Windows Server. *learn.microsoft.com*. URL: https://learn.microsoft.com/en-us/troubleshoot/windows-server/networking/default-dynamic-port-range-tcpip-chang

20. Umberto Michelucci. An Introduction to Autoencoders. *arXiv (Cornell University)*, Jan. 2022, doi: https://doi.org/10.48550/arxiv.2201.03898

21. Huber P. J. Robust Estimation of a Location Parameter. Springer series in statistics, pp. 492–518, Jan. 1992, doi: https://doi.org/10.1007/978-1-4612-4380-9_35.