https://doi.org/10.31891/2219-9365-2024-79-29
УДК 004.8, 004.056

ВОЛОКИТА Артем
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
https://orcid.org/0000-0001-9069-5544
e-mail: artem.volokita@kpi.ua
МЕЛЕНЧУКОВ Микита
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»
https://orcid.org/0009-0005-6615-4306
e-mail: melenchukov.nikita@gmail.com

ДОСЛІДЖЕННЯ МОДЕЛЕЙ ВИЯВЛЕННЯ АТАК НА РОЗПОДІЛЕНІ СИСТЕМИ ЗА ДОПОМОГОЮ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

Актуальною є задача захисту розподілених систем обробки даних, які є вразливими до кібератак, за допомогою спеціалізованих систем виявлення атак (СВА). У цій роботі досліджено застосування нейронних мереж для виявлення атак на мережі із розподіленою архітектурою. Проведено аналіз та попередню обробку даних для машинного навчання. Розроблено групу моделей виявлення атак на розподілені системи. Виконано експериментальні дослідження для знаходження оптимального рішення з найвищою точністю визначення атак на такі системи.
Ключові слова: машинне навчання, штучні нейронні мережі, розподілені системи.

VOLOKYTA Artem, MELENCHUKOV Mykyta
National Technical University of Ukraine “Ihor Sikorskyi Kyiv Polytechnic Institute"

RESEARCH ON ATTACK DETECTION MODELS FOR DISTRIBUTED SYSTEMS USING CONVOLUTIONAL NEURAL NETWORKS

With the exponential growth of information that is available to mankind, the corresponding need to process the growing volumes of data also grows. Distributed systems are great for solving such problems, but in turn have certain disadvantages, among them - susceptibility to interference attacks. Such attacks can be prevented by detecting them in time and taking appropriate actions necessary for protection. This article examines the approaches and methods by which a good level of attack detection can be achieved, to test the hypotheses, experiments were performed on self-developed models of convolutional neural networks.

The available articles on this topic were reviewed and a conclusion was drawn about the rapid development of this direction and the need to continue it.

By combining experiments with different models of neural networks, a variety of processing of input data was performed in order to increase the accuracy and quality of intervention detection.The results of the conducted experiments and the corresponding preparations for them are analyzed in detail. As a result, this article provides important information about effective methods and approaches to improve attack detection accuracy and distributed system security, respectively.

Further directions for the development of this topic are also given, which are interesting and important for new research
Keywords: machine learning, artificial neural networks, distributed systems.

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ’ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

На сьогоднішній день обсяги даних що доступні людству постійно зростають, як наслідок існує необхідність обробляти дедалі більше інформації. До вирішення задач такого типу добре пристосовані розподілені системи. Через свою популярність, такі системи стають мішенями для атак, кількість видів яких теж постійно збільшується. У цій статті розроблено декілька моделей нейронних мереж що дозволяють виявити такі втручання, проаналізовано дані які необхідні для навчання мережі, зроблена їх попередня обробка, досліджено вплив параметрів моделі на точність детекції.

Розподілені системи та атаки на них знаходяться у постійному розвитку. Існуючі підходи створення баз даних та умов для виявлення атак потребують постійного оновлення для підтримки ефективності захисту. Детекція втручання за допомогою нейронних мереж, у свою чергу, дозволяє швидко адаптуватись до нових підходів нападу. Разом із цим розподілені системи містять велику кількість різноманітної інформації яку можна аналізувати – логи, об’єм трафіку, аналіз пакетів, послідовність дій. Також існує велика кількість видів нейронних мереж(НМ). Як наслідок виникає необхідність пошуку ключових принципів побудови оптимальних моделей НМ та кращих підходів до обробки даних які можуть вказати на наявність нападу.

АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Системам виявлення атак (СВА) присвячено багато статей, цей напрям зараз активно розвивається[1, 2]. Автори [3] наголошують що кібербезпека стала критичною проблемою у останні роки у зв’язку із поширенням інформаційних технологій. Наслідком цього є зростання популярності СВА, та

зміщення розвитку СВА у сторону застосування машинного навчання. Цей напрямок привертає значну увагу дослідників. У статті автори розробили систему Passban яка запобігає втручанню у систему пристроїв Інтернету речей за допомогою розгортання вузлів СВА на недорогих шлюзах. Недоліком такого підходу є невміння системи самостійно адаптуватись до нових атак та необхідність постійних оновлень. Іншим викликом для такої системи є проблема масштабування, а саме зростаючого об'єму даних при розширенні Інтернету речей та взаємодія між шлюзами.

У статті [4] автори представляють розробку СВА що працює на основі глибокого трансферного навчання і поєднує у собі згорткові нейронні мережі, генетичні алгоритми, та поєднання декількох моделей шляхом голосування. Навчання відбувається на базі датасету Інтернету речей Edge_IIoTset, у попередній обробці автори виконують трансформацію вхідних даних у зображення з метою подальшої обробки у згортковій нейронній мережі. За допомогою генетичного алгоритму виконується пошук оптимальних гіперпараметрів декількох моделей, результати детекції яких потім поєднуються шляхом голосування. Такий підхід дозволив авторам досягти точності виявлення у 100% для 14 різних видів атак.

За допомогою поєднання трансферного навчання і згорткової нейронної мережі із короткою пам'яттю автори статті [5] створили власну модель СВА IDS-INT що була протестована на декількох датасетах – UNSW-NB15, CIC-IDS2017 і вже класичний NSL-KDD.

Інший підхід до виявлення втручань у системи Інтернету речей застосували автори статті [6]. Використовуючи алгоритми дерево рішень, випадковий ліс(RF), k найближчих сусідів та метод опорних векторів, вони отримали власні моделі СВА. Для навчання застосовувався один із найновіших доступних датасетів систем Інтернету речей – IoTID20. Крім цього автори виконали попередню обробку даних, за допомогою алгоритму відбору ознак на основі кореляції. Автори статті [7] створили систему виявлення втручань із застосуванням машинного навчання, різних алгоритмів, наприклад, випадкового лісу. Вони теж працювали із системою інтернет речей, та використовували датасети TON-IoT і UNSW-NB15. В результаті вдалось досягти точності детекції у 96.04%-100%. У статті [8] автори працюють над схожою задачею але залучають до нейронної мережі LSTM (довгу короткочасну пам'ять), це дозволило отримати точність у 97%-100%.

У дослідженні [9] автори створили свою СВА за допомогою використання техніки навчання без нагляду. Творці такої системи прагнули створити рішення що може ефективно працювати в умовах обмеженої апаратної потужності. За допомогою навчання без нагляду, запропонована система може тренуватись та адаптуватись до змін мережевої поведінки не вимагаючи постійних оновлень. Утворену модель автори назвали Kitsune.

ВИДІЛЕННЯ НЕДОСЛІДЖЕНИХ ЧАСТИН ЗАГАЛЬНОЇ ПРОБЛЕМИ

Враховуючи розглянуті дослідження і публікації зроблено висновок що напрямок виявлення атак на розподілені системи знаходиться у активному розвитку разом із СВА. Нові підходи до виявлення залучають різні алгоритми машинного навчання, а їх автори виконують різні попередні обробки та маніпуляції із вхідними даними з метою підвищення точності (Accuracy), швидкодії, стійкості, ефективності. Таким чином існує потреба у подальшому розвитку СВА, створенню нових моделей та ключових підходів що дозволяють досягнути кращих результатів.

ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є створення нової моделі системи виявлення атак та дослідження оптимальних підходів для підвищення точності та ефективності виявлення атак на розподілену систему за допомогою згорткових нейронних мереж на прикладі датасету CICIDS2018.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Система виявлення атак(СВА) – це така програмна чи апаратна система, що націлена на виявлення кібератак з метою підтримки безпеки комп'ютерної системи [10]. Детекція втручання відбувається завдяки аналізу різних даних – логів, інтернет-пакетів, об'єму трафіку. В залежності від виду атаки яку намагаються знайти, певні дані системи підходять краще ніж інші. У цій статті було виконано розробку системи яка розрізняє доброякісний трафік та такий що містить Bruteforce-SSH чи Bruteforce-FTP атаки. Bruteforce атака це таке втручання у систему, при якому атакуючий застосовує різноманітні автоматизуючі скрипти чи програми які дозволяють перебрати усі можливі комбінації ключів для авторизації, з метою отримання доступу до системи. Відповідно мета Bruteforce-SSH втручання це зламати пароль для отримання вхідних прав SSH, а Bruteforce-FTP – для отримання доступу до FTP-сервера.

Для вирішення поставленої задачі використовувався один із найновіших СВА-датасетів CICIDS2018. Він був створений у 2018-ому році в університеті Нью-Брансуика для аналізу даних DDoS атак, та складається із логів серверів університету, які містять чисельну кількість DDoS атак що були проведені протягом спеціально дозволеного періоду, для збору інформації про кібератаки. Кожного дня відбувались різні категорії атак, у цій роботі використовуються записи що були отримані 14.02.2018 і

містять у собі Bruteforce-SSH, Bruteforce-FTP та доброякісний трафік. Усього доступно 1048575 записів, кожен запис містить 80 полів – загалом 640MB. Не всі з наявних полів заповнені, а деякі колонки не несуть у собі корисної інформації. Тож було відкинуто записи із нульовими полями та такі колонки як "Timestamp", "Protocol","PSH Flag Cnt","Init Fwd Win Byts","Flow Byts/s","Flow Pkts/s" – вони несли у собі інформацію про час з'єднання, протокол, кількість PSH флагів(флаг негайності), початковий розмір вікна передачі, об'єми трафіку. Записів які помічені як доброякісні і такі що містять Bruteforce-FTP чи Bruteforce-SSH атаки, різна кількість (рисунок 1).

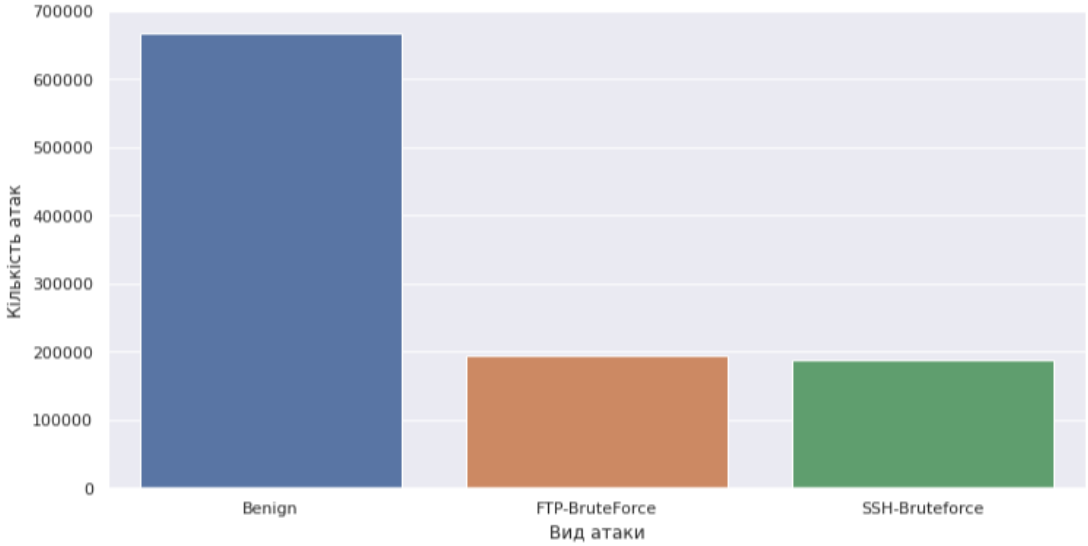

Рис. 1. Розподіл кількості атак за типом

Джерело: розроблено авторами

Для уникнення упередженості вхідних даних, було відібрано по 20 000 рядків кожного типу та складено новий датасет із 60 000 записів (рис. 2).

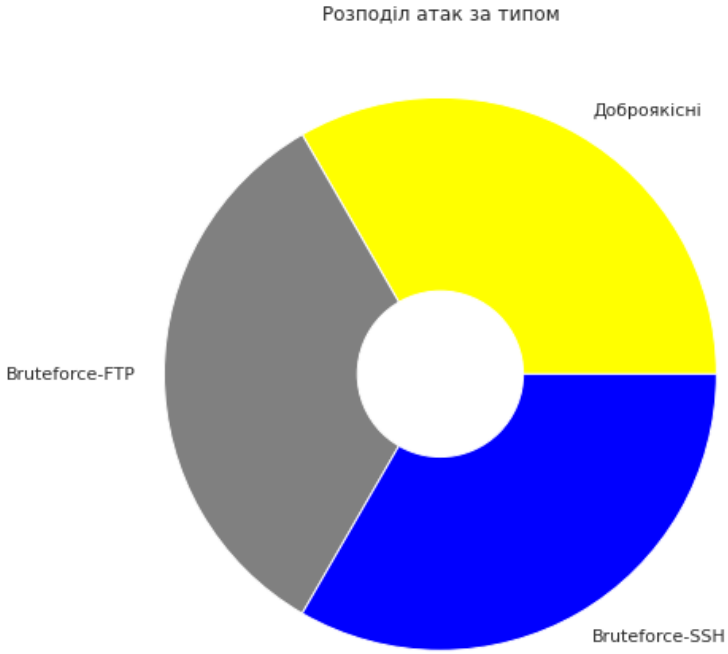

Рис. 2. Розподіл атак за типом після відбору даних для уникнення їх упередженості

Джерело: розроблено авторами

Для розпізнавання втручань було розроблено 9 моделей нейромереж що різняться функцією активації та оптимізатором. Застосовувались такі функції активації як ReLU, SiLU, sigmoid. ReLU(recified linear unit) це популярна для застосування функція активації що використовується у багатьох моделях

нейромереж. Перевагою цієї функції у порівнянні із альтернативами є уникнення проблеми занепадаючих градієнтів та її швидкість обчислення. Ця функція описується наступною формулою:

$$f(x) = \max(0, x)$$

Формула 1 Функція ReLU.

Наступною функцією яка була обрана для дослідження є SiLU(sigmoid linear unit), у порівнянні із ReLU вона є більш вимогливою до обчислень, але не вразлива до проблеми мертвих нейронів, уникає затухання градієнтів що дозволяє пришвидшити навчання моделі, та зробити цей процес більш стабільним. Описується ця функція наступною формулою:

$$f(x) = x * \sigma(x)$$

Де $\sigma(x) = \frac{1}{1 + e^{-x}}$

Де λ та α є константами-гіперпараметрами.

Формула 2 Функція SiLU.

Останньою із досліджуваних функцій активаціє було обрано класичну для нейромереж функцію sigmoid. Від попередніх опцій вона відрізняється обмеженістю вихідних даних у рамках діапазону [0, 1], а також є пристосованою до різноманітних задач бінарної класифікації де необхідно ймовірнісно оцінити результат. Задається вона наступним чином:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Формула 3 Функція sigmoid.

Іншим гіперпараметром що досліджувався був тип оптимізатору – алгоритму для налаштування ваг нейронної мережі під час навчання. Основна задача оптимізатора – мінімізувати функцію втрат. У цій статті було обрано такі оптимізатори як Adagrad, Adam і Nadam. Оптимізатор Adagrad адаптує швидкість навчання для кожного параметра окремо, балансуючи її в залежності від градієнтів. Цей варіант краще підходить для розріджених даних, але швидкість навчання може почати страждати через певний час застосування, викликаючи проблему застрягання у локальних мінімумах. Adam вирішує цю проблему завдяки зберіганню середнього та середньоквадратичного значення градієнтів для кожного параметра. Оптимізатор Nadam поєднує у собі оптимізатор Adam та Nesterov Accelerated Gradient для обчислення градієнтів після прогнозування майбутнього положення. Це викликає певні складності у налаштуванні. Але робить процес оптимізації ще стабільнішим та швидшим.

Для виявлення атак на розподілену систему по логам із датасету CICIDS2018 було використано згорткову нейронну мережу архітектура якої відображена на рисунку 3.

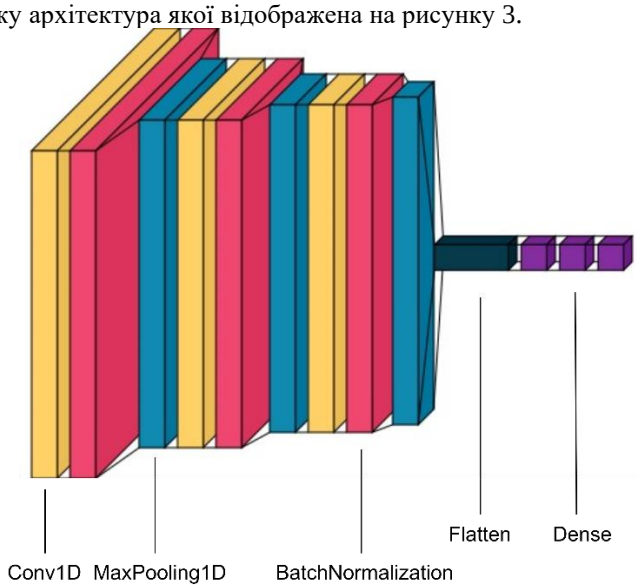

Рис. 3. Модель запропонованої згорткової мережі

Джерело: розроблено авторами

Сама мережа має 72 вхідних нейрони та складається із 5 видів різних шарів: Conv1D – обчислює скалярний добуток між вагами і вхідним сигналом. BatchNormaliztion – нормалізує виходи попередніх шарів. MaxPooling1D – зменшує розмірність даних завдяки пошуку максимального значення у підгрупі вхідних значень. Flatten (шар випрямлення) перетворює багатовимірний тензор у одновимірний вектор. Dense (повнозв'язний шар) виконує лінійне перетворення вхідних даних. Функція активації задавалась на шарах Conv1D та Dense, оптимізатор обирався для всієї моделі.

Навчання виконувалось у 30 ітерацій, та дозволило отримати такі результати (Таблиця 1):

Таблиця 1

Точність виявлення атак відповідно до обраної функції активації та оптимізатора

	Adagrad	Adam	Nadam
ReLU	97.883%	33.75%	83.316%
SiLU	99.966%	61.733%	99.916%
Sigmoid	99.966%	100%	100%

Джерело: розроблено авторами

Точність у динаміці(під час навчання), змінювалась для кожного набору гіперпараметрів по різному, графік цих змін відображений на рисунку 4.

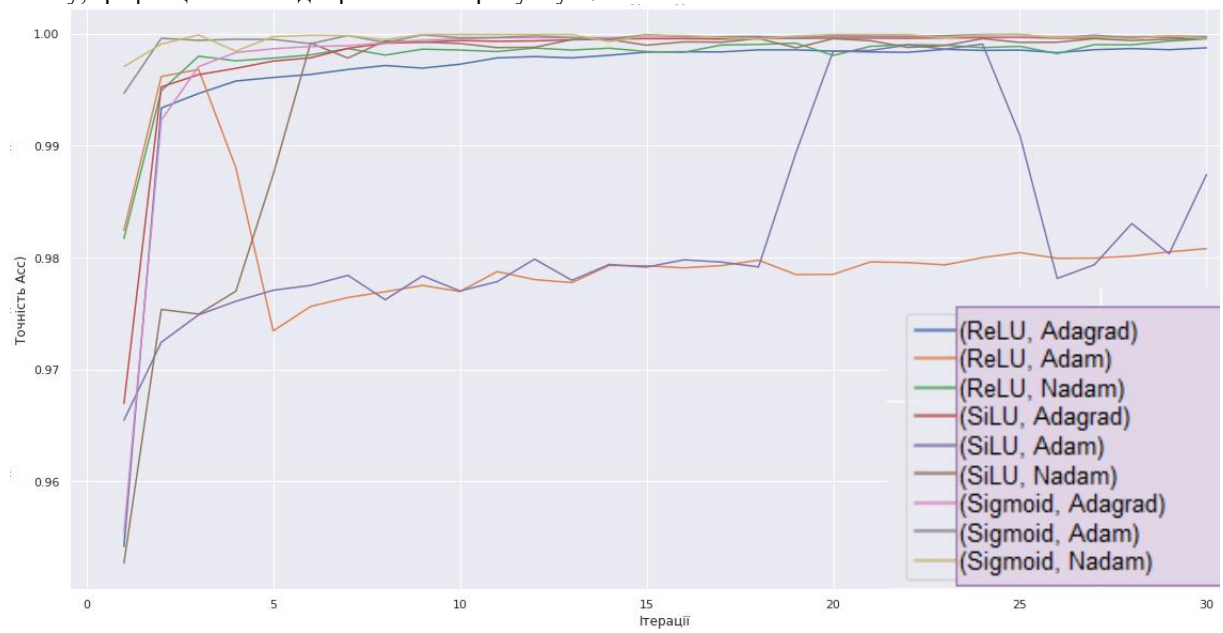

Рис. 4. Зміна точності виявлення кібератак

Джерело: розроблено авторами

Варто відмітити значні коливання у точності для комбінації (relu, adam), (silu, adam), (silu, nadam).

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

В ході виконаного проєкту досліджено багато гіпотез і підходів. Вхідний датасет був створений в реальних умовах проведення кібератак, тож не дивно що він містив пропуски та пусті поля. Таким чином його попередня обробка була вимушеною. Крім цього була помітна нерівність у кількості даних між різними типами втручань у систему (рисунок 1), тож було прийнято рішення відібрати однакову кількість записів для кожного виду кібератаки відповідно. Деякі поля, такі як час з'єднання чи об'єм трафіку, не були релевантні у досліджуваному варіанті атак, тож були відкинуті. Завдяки різним варіантам оптимізаторів так функцій активації, було утворено 9 різним згорткових нейронних мереж. Оптимізатор Adagrad стабільно показував гарні результати для всіх функцій активації, та все ж кращою комбінацією виявились пари Adam, sigmoid і Nadam, sigmoid – завдяки такому поєднанню було досягнуто відмінний результат у 100% точності. Що, враховуючи попередні показники оптимізатору Adam для ReLU і SiLU у 33.75% та 61.733%, є неочікуваним результатом.

У подальшому це дослідження можна розвинути, залучивши нові дані із більшою кількістю параметрів, та іншими обмеженнями, наприклад малим об'ємом інформації, чи вибірок із малою кількістю записів що відносяться до кібератак. Це вимусить шукати інші шляхи до балансування вхідних даних для пошуку способу уникнення упередженості моделі. Крім цього варто дослідити обробку вхідних даних для

створення можливості роботи із нейронними мережами що мають багато вхідних нейронів, наприклад U-net чи VGG16.

Література

1. Liao, H., Lin, C. R., Lin, Y., & Tung, K. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, *36*(1), 16–24. DOI:10.1016/j.jnca.2012.09.004

2. Zarpelão, B. B., Miani, R. S., Kawakani, C. T., & De Alvarenga, S. C. (2017). A survey of intrusion detection in Internet of Things. *Journal of Network and Computer Applications*, *84*, 25–37. DOI:10.1016/j.jnca.2017.02.009

3. Eskandari, M., Janjua, Z. H., Vecchio, M., & Antonelli, F. (2020). PASSBAN IDS: An intelligent Anomaly-Based intrusion Detection System for IoT edge devices. *IEEE Internet of Things Journal*, *7*(8), 6882–6897. DOI:10.1109/jiot.2020.2970501

4. Latif, S., Boulila, W., Koubaa, A., Zou, Z., & Ahmad, J. (2024). DTL-IDS: An optimized Intrusion Detection Framework using Deep Transfer Learning and Genetic Algorithm. *Journal of Network and Computer Applications*, *221*, 103784. DOI:10.1016/j.jnca.2023.103784

5. Ullah, F., Ullah, S., Srivastava, G., & Lin, J. C. (2023). IDS-INT: Intrusion detection system using transformer-based transfer learning for imbalanced network traffic. *Digital Communications and Networks*. DOI:10.1016/j.dcan.2023.03.008

6. Altulaihan, E., Almaiah, M. A., & Aljughaiman, A. (2024). Anomaly Detection IDS for detecting DOS attacks in IoT networks based on machine learning algorithms. *Sensors*, *24*(2), 713. DOI:10.3390/s24020713

7. Al-Ambusaidi, M., Yinjun, Z., Muhammad, Y., & Yahya, A. (2023). ML-IDS: an efficient ML-enabled intrusion detection system for securing IoT networks and applications. *Soft Computing*, *28*(2), 1765–1784. DOI:10.1007/s00500-023-09452-7

8. Otoum, Y., Liu, D., & Nayak, A. (2019). DL-IDS: a deep learning–based intrusion detection framework for securing IoT. *Transactions on Emerging Telecommunications Technologies*, *33*(3). DOI:10.1002/ett.3803

9. Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). KITSUNE: an ensemble of autoencoders for online network intrusion detection. *arXiv (Cornell University)*. DOI:10.48550/arxiv.1802.09089

10. Volokyta, A., & Melenchukov, M. (2024). Neural networks in detecting attacks on distributed systems. Technical sciences and technologies, *1(35)*, 135–145. DOI:10.25140/2411-5363-2024-1(35)-135-145

References

1. Liao, H., Lin, C. R., Lin, Y., & Tung, K. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, *36*(1), 16–24. DOI:10.1016/j.jnca.2012.09.004

2. Zarpelão, B. B., Miani, R. S., Kawakani, C. T., & De Alvarenga, S. C. (2017). A survey of intrusion detection in Internet of Things. *Journal of Network and Computer Applications*, *84*, 25–37. DOI:10.1016/j.jnca.2017.02.009

3. Eskandari, M., Janjua, Z. H., Vecchio, M., & Antonelli, F. (2020). PASSBAN IDS: An intelligent Anomaly-Based intrusion Detection System for IoT edge devices. *IEEE Internet of Things Journal*, *7*(8), 6882–6897. DOI:10.1109/jiot.2020.2970501

4. Latif, S., Boulila, W., Koubaa, A., Zou, Z., & Ahmad, J. (2024). DTL-IDS: An optimized Intrusion Detection Framework using Deep Transfer Learning and Genetic Algorithm. *Journal of Network and Computer Applications*, *221*, 103784. DOI:10.1016/j.jnca.2023.103784

5. Ullah, F., Ullah, S., Srivastava, G., & Lin, J. C. (2023). IDS-INT: Intrusion detection system using transformer-based transfer learning for imbalanced network traffic. *Digital Communications and Networks*. DOI:10.1016/j.dcan.2023.03.008

6. Altulaihan, E., Almaiah, M. A., & Aljughaiman, A. (2024). Anomaly Detection IDS for detecting DOS attacks in IoT networks based on machine learning algorithms. *Sensors*, *24*(2), 713. DOI:10.3390/s24020713

7. Al-Ambusaidi, M., Yinjun, Z., Muhammad, Y., & Yahya, A. (2023). ML-IDS: an efficient ML-enabled intrusion detection system for securing IoT networks and applications. *Soft Computing*, *28*(2), 1765–1784. DOI:10.1007/s00500-023-09452-7

8. Otoum, Y., Liu, D., & Nayak, A. (2019). DL-IDS: a deep learning–based intrusion detection framework for securing IoT. *Transactions on Emerging Telecommunications Technologies*, *33*(3). DOI:10.1002/ett.3803

9. Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). KITSUNE: an ensemble of autoencoders for online network intrusion detection. *arXiv (Cornell University)*. DOI:10.48550/arxiv.1802.09089

10. Volokyta, A., & Melenchukov, M. (2024). Neural networks in detecting attacks on distributed systems. Technical sciences and technologies, *1(35)*, 135–145. DOI:10.25140/2411-5363-2024-1(35)-135-145