ІНФОРМАТИКА, ОБЧИСЛЮВАЛЬНА ТЕХНІКА ТА АВТОМАТИЗАЦІЯ

УДК 004.056.5:004.7
DOI https://doi.org/10.32838/TNU-2663-5941/2020.3-1/12

Бабкін А.А.
Запорізький національний університет

Кудін О.В.
Запорізький національний університет

ОГЛЯД НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ СИСТЕМ ВИЯВЛЕННЯ ВТОРГНЕНЬ

На сучасному етапі розвитку інформаційного суспільства особливо актуальним завданням є забезпечення інформаційної безпеки персональних і корпоративних даних. Захист даних в інтернеті речей і вбудованих системах є темою публікацій останніх років. Одним із аспектів захисту інформації є розроблення спеціалізованого програмного забезпечення, завдання якого полягає у виявленні потенційних загроз із подальшим автоматичним реагуванням або інформуванням користувачів. Прикладом таких програмних засобів є системи виявлення вторгнень, що використовуються для забезпечення захисту комп'ютерних мереж від несанкціонованого доступу. Нині можна виділити такі основні напрями в розробленні подібних систем: підходи на основі певного набору правил; методи автоматизованого виявлення аномалій трафіку. Передбачається, що поведінка зловмисника в комп'ютерній мережі відрізняється від цифрових слідів, які залишає звичайний користувач. Отже, зазвичай завдання виявлення вторгнень зводиться до аналізу мережевого трафіку й розроблення методів розпізнавання. Одним із потужним засобів автоматичного пошуку шаблонів даних є нейронні мережі, які знаходять застосування й у системах виявлення атак. Особливостями роботи таких систем є висока швидкість навчання та адаптивність до нових типів атак. Зазвичай таке програмне забезпечення захисту даних працює в режимі реального часу й аналізує трафік як усередині мережі, так і зовнішні запити. Стаття присвячена аналізу останніх публікацій із побудови нейромережевих моделей у цій галузі. Акцент робиться на роботах, які пропонують методи побудови гібридних систем на основі поєднання нейронних мереж з іншими методами машинного навчання. Саме такі підходи отримали значні результати в комп'ютерній галузі та у сфері машинного перекладу. На основі аналізу попередніх публікацій пропонуються шляхи розвитку систем виявлення атак.

Ключові слова: *система виявлення вторгнень, нейронна мережа, захист інформації, гібридна система, ансаблеве навчання.*

Постановка проблеми. Забезпечення захисту інформації є комплексною проблемою, яка включає не тільки технічний аспект, а й правовий, соціальний, культурний тощо. Однак саме розвиток апаратних і програмних засобів несанкціонованого доступу потребує відповідного адаптування технічного інструментарію захисту даних. Так звані системи виявлення вторгнень (англ. Intrusion Detection System, IDS) є досить поширеним засобом захисту інформації користувачів і підприємств. Головною метою таких систем є ідентифікація та фільтрація потенційно шкідливих запитів (атак) у комп'ютерних мережах [1]. Математична модель цього процесу зводиться до добре відомого завдання розпізнавання образів [2], а саме класифікації або кластеризації даних з метою виявлення аномальних даних у трафіку комп'ютерної мережі. Останнім часом для розв'язання завдання розпізнавання широко застосовуються нейронні мережі [3]. Це пояснюється потенційною здатністю нейромережевих моделей до генералізації, тобто виявлення нових типів атак, однак остаточно це завдання досі не вирішено. Схема застосування нейронних мереж у завданнях захисту даних не відрізняється від загальноприйнятної [3]. До недоліків цього підходу можна зарахувати

необхідність досить великого набору даних для ефективного навчання та вимогливість до ресурсів обчислювальних систем. При цьому висока ефективність подібних систем на основі методів машинного навчання не гарантується. Отже, актуальним завданням є аналітичний огляд наявних публікацій із застосування нейронних мереж у системах виявлення вторгнень з метою виявлення можливих напрямів розвитку технології.

Аналіз останніх досліджень і публікацій. В оглядових статтях [4–10] наводиться класифікація найбільш поширених методів роботи IDS: підходи на основі виявлення сигнатур потенційних атак і на основі пошуку аномалій у даних. У роботі [6] наводиться детальний аналіз переваг і недоліків цих двох методів. Загальним є припущення, що сліди діяльності зловмисника в комп'ютерній мережі відрізняються від стандартної поведінки звичайного користувача.

У роботі [4] проведено класифікацію наявних методів глибинного навчання на генеративні, описові та гібридні. Генеративні методи, наприклад, нейронні мережі автокодувальники, використовуються для абстрактного представлення ознак мережевого запиту або розширення навчальної вибірки. Описові моделі використовуються зазвичай під час навчання з учителем, тобто коли є набір даних, у якому розмічені зловмисні та звичайні запити. Основним інструментом тут є згорткові нейронні мережі та багатошаровий перцептрон. Аналогічний огляд із посиланням на велику кількість статей із різних завдань інформаційної безпеки наведено в роботі [7].

Огляд наявних методів колективного інтелекту (англ. Swarm Intelligence, SI), які полягають у поєднанні великої кількості інтелектуальних агентів у розподілену систему для розв'язання завдань оптимізації, наводиться в праці [5]. Показано, що агенти можуть використовуватися для пошуку правил класифікації мережевих запитів або аномалій, а також наводиться огляд робіт із моделювання гібридних систем колективного інтелекту та штучних нейронних мереж.

Детальна таксономія підходів до розроблення IDS на базі методів машинного навчання наводиться в роботах [6; 9; 10]. Як ознаки для класифікації виділяють у тому числі й джерело даних, яке використовується для виявлення вторгнень: обладнання комп'ютерної мережі або лог-файли користувацької операційної системи. Також наводяться посилання на роботи з використання більшості з відомих алгоритмів машинного навчання, у тому числі різних типів нейронних мереж.

Статті [8; 9] містять огляд публікацій із застосування генетичних алгоритмів і нечіткої логіки під час розроблення систем виявлення атак. Перевагою генетичних алгоритмів є можливість автоматичного пошуку оптимальної комбінації параметрів алгоритму класифікації, наприклад, нейронної мережі, що є важливим етапом розв'язання завдання. Класифікатор із використанням методів нечіткої логіки може генерувати набір правил, що, у свою чергу, можуть аналізуватися спеціалістами з інформаційної безпеки.

Загальною рисою наведених оглядів є досить детальна класифікація методів машинного навчання та посилання на роботи з їх застосування в завданнях виявлення атак. Значно менше уваги приділяється огляду робіт, які поєднують декілька алгоритмів за допомогою методів побудови ансамблів або інших гібридних моделей.

Постановка завдання. Стаття присвячена аналізу останніх публікацій із застосування нейронних мереж у розробленні програмних систем виявлення вторгнень. Особлива увага приділяється аналізу робіт з ансамблевого навчання та побудови гібридних систем у цій галузі. На основі проведеного аналізу пропонуються можливі шляхи розвитку.

Виклад основного матеріалу дослідження. Розроблення програмного забезпечення для запобігання несанкціонованому доступу та захисту даних в інформаційних системах є актуальним завданням сучасної інформатики. Труднощі в розробленні таких програмних засобів пов'язані з необхідністю обробки великої кількості даних про мережеві з'єднання в режимі реального часу, а також із вимогою адаптованості таких систем до нових типів атак (атаки нульового дня).

Нині в науковій літературі запропоновано загальні підходи до розроблення та оцінювання параметрів безпеки нейромережевих моделей захисту інформації [11; 12; 13]. Огляд стандартів мережевої безпеки наведено в праці [14].

Досить велику кількість вітчизняних і зарубіжних робіт присвячено застосуванню нейронних мереж у завданнях виявлення вторгнень [15–37], класифікації шкідливого програмного забезпечення [38] тощо. Особливістю цих робіт є застосування методів машинного навчання та архітектур штучних нейронних мереж, які вже стали класичними в таких застосуваннях, як комп'ютерний зір.

У роботі [15] пропонується використання згорткових нейронних мереж з одномірними шарами для побудови класифікатора мережевих запитів. Виявлено, що на відкритих наборах

даних така модель демонструє вищу точність, ніж моделі на основі багатошарового перцептрону чи рекурентних мереж.

Структурно-логічні моделі виявлення шкідливого програмного забезпечення розглядаються в роботі [16], де запропоновано сигнатурно-кубітні методи синтезу логічних схем.

У праці [17] пропонується архітектура розподіленої хмарної системи виявлення для розумних міст (англ. Smart City), використовуються глибинні нейронні мережі довіри (англ. Deep Belief Network, DBN) і дерева рішень.

У великій кількість робіт застосовуються нейронні мережі-автокодувальники. Наприклад у праці [18] автокодувальник використовується для вилучення інформативних ознак, а як класифікатор виступає багатошаровий перцептрон.

Як і в завданнях комп'ютерного зору, важливим питанням є підготовка наборів даних для тренування нейромережевих моделей. Робота [19] пропонує рекурсивний метод знаходження кращого набору ознак, які можна використовувати для класифікації мережевих запитів.

Методи нечіткої логіки в поєднанні з нейронними мережами утворюють окремий клас класифікаторів – нечіткі нейромережеві моделі. У праці [20] розроблено підхід до класифікації на базі нечітких правил. Завдяки цьому підходу можна отримати набір умов у форматі «Якщо... то...», які є підґрунтям для прийняття рішення нейронною мережею.

У роботах [21; 23] використовуються генеративні змагальні нейронні мережі (англ. Generative adversarial networks, GANs). Систему на основі напівавтоматичного навчання (англ. Semi-supervised learning) із застосуванням генеративних нейронних мереж запропоновано в праці [21]. Ідея підходу полягає в генеруванні додаткових даних, які імітують запити вторгнення. Розподілену систему виявлення атак у мережах інтернету речей (англ. Internet of Things, IoT) розроблено авторами роботи [23]. Особливістю є відсутність необхідності мати спільну базу даних для роботи вузлів системи.

Системи виявлення вторгнень знаходять застосунок не тільки в стандартних комп'ютерних мережах. Так, у праці [24] запропоновано архітектуру рекурентних нейронних мереж для виявлення вторгнень в електронну мережу автомобілів.

Авторами статті [25] розробляється методологія отримання інформації з розрізнених баз даних про проведені атаки. Пропонується метод знаходження аномалій і розв'язується обернене завдання – визначаються ознаки аномальних запитів. Такий підхід дає експертам змогу отримувати зрозумілий опис вторгнень для подальшого аналізу.

Роботи [26; 30; 31; 33] присвячено аналізу трафіку в комп'ютерних мережах індустріальних підприємств. Випадкові дерева та метод опорних векторів використовуються для аналізу трафіку в мережах автоматизованих систем керування виробництвом [26]. Автори роботи [30] для тієї ж мети застосовують рекурентні нейромоделі. У праці [31] описано сім різних алгоритмів класифікації та особливості протоколів індустріальних мереж.

Ще одним прикладом використання підходів із галузі комп'ютерного зору є використання гібридних систем на базі згорткових мереж у поєднанні з рекурентною архітектурою. У роботі [27] подібна система використовується для багатокласової класифікації.

У роботі [28] для класифікації запитів застосовується глибинна нейронна прямого мережа та метод аналізу головних компонент під час визначення найбільш інформативних ознак. Отже, простір ознак, що використовуються класифікатором, зменшується. Унаслідок цього система може швидше оброблювати дані, що під час аналізу трафіку в реальному часі є суттєвою перевагою.

Також відомим підходом до об'єднання декількох моделей машинного навчання є ітеративне покращення прогнозу на деяких підмножинах вихідних даних. У праці [29] використовується подібна каскадна система двох автокодувальників. Ці методи забезпечують перетворення та стиснення вхідних сигналів.

Рекурентні нейронні мережі застосовано в роботах [32; 34]. Зокрема, автори [34] пропонують гібридну систему, особливістю якої є здатність класифікувати необроблені дані мережевого трафіку.

У роботі [35] протестовано застосунок KNIME, за допомогою якого виконано класифікацію різними алгоритмами машинного навчання та проведено порівняння.

Набагато менше робіт присвячено питанню тестування стійкості нейромережевих систем виявлення вторгнень до зовнішніх атак. Наприклад, у роботі [37] розглядається випадок змагальних прикладів (англ. Adversarial examples), коли зловмисник намагається переналаштувати нейромережеву систему виявлення атак за допомогою штучно згенерованих оманних прикладів даних.

Досить часто під час використання нейромережевих моделей, наприклад, у завданнях розпізнавання, автоматичного перекладу, застосовуються різні підходи до поєднання відповідей класифікаторів декількох типів у єдиний результат. Такі підходи мають загальну назву «ансамблеве навчання».

Особливості ансамблювання різних алгоритмів у завданнях виявлення атак розглядаються в роботах [36; 39]. Так, згорткові нейронні мережі в поєднанні з рекурентною та глибинною архітектурою використовуються в праці [36]. Результати цих класифікаторів поєднуються в спільну відповідь за допомогою алгоритму лісу випадкових дерев рішень.

Висновки. Отже, особливістю всіх розглянутих вище робіт останніх років є тенденція до побудови гібридних систем та ансамблів, заснованих на поєднанні декількох типів нейронних мереж і/або інших методів машинного навчання. Більшість підходів базується на навчанні з учителем (завдання класифікації) або без учителя (завдання кластеризації). Також велика увага приділяється попередній підготовці даних для навчання. Тут частіше використовуються автокодувальники або метод головних компонент для зменшення простору ознак і прискорення наступних обчислень.

Перевагами попередніх оглядових робіт є досить детальний опис особливостей систем виявлення вторгнень і базових методів машинного навчання. Водночас майже не приділялася увага методам ансамблювання та побудови гібридних систем, які активно застосовуються в інших застосунках нейронних мереж.

Особливістю цього огляду є акцент на публікаціях із поєднання декількох підходів до побудови систем виявлення мережевих атак з метою покращення результуючої точності.

Перспективи подальшого розвитку в цій галузі можуть бути пов'язані з подальшим розвитком гібридних систем і використанням таких підходів, як метанавчання [40], що може зменшити час тренування подібних систем і підвищити здатність до генералізації.

Список літератури:

1. Javaid A.Y., Niyaz Q., Sun W., Alam M. A Deep Learning Approach for Network Intrusion Detection System. 2016. URL: https://doi.org/10.4108/eai.3-12-2015.2262516.

2. Мазуров В.Д. Математические методы распознавания образов : учебное пособие. 2-е изд. Екатеринбург : Урал. ун-т, 2010. 101 с.

3. Geron A. Hands-On Machine Learning with Scikit-Learn and TensorFlow. Sebastopol : O`Reilly, 2017. 861 p.

4. Azawii A., Al-Janabi S.F.T., Al-Khateeb B. Survey on Intrusion Detection Systems based on Deep Learning. *Periodicals of Engineering and Natural Sciences*. 2019. Vol. 7 (3). P. 1074–1095. URL: http://dx.doi.org/10.21533/pen.v7i3.635.

5. Kolias C., Kambourakis G., Maragoudakis M. Swarm intelligence in intrusion detection: A survey. *Computers & Security The International Source of Innovation for the Information Security and IT Audit Professional.* 2011. P. 625–642. URL: https://doi.org/10.1016/j.cose.2011.08.009.

6. Liu. H, Lang B. Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey. *Appl. Sci*. 2019. Vol. 9 (20). 4396 p. URL: https://doi.org/10.3390/app9204396.

7. Berman D.S., Buczak A.L., Chavis J.S., Corbett C.L. A Survey of Deep Learning Methods for Cyber Security. *Information*. 2019. № 10. 122 p. URL: https://doi.org/10.3390/info10040122.

8. Zamani M., Movahedi M. Machine Learning Techniques for Intrusion Detection. *Arxiv*. 2015. URL: https://arxiv.org/abs/1312.2177.

9. Hodo E., Bellekens X., Hamilton A., Tachtatzis C., Atkinson R. Shallow and Deep Networks Intrusion Detection System: A Taxonomy and Survey. *Arxiv*. 2017. URL: https://arxiv.org/abs/1701.02145.

10. Pande S.D., Khamparia A. A Review on Detection of DDOS Attack Using Machine Learning and Deep Learning Techniques. *Think India journal.* 2019. Vol. 22 (16). P. 2035–2043.

11. Корченко О., Терейковський І., Білощинський А. Методологія розроблення нейромережевих засобів інформаційної безпеки інтернет-орієнтованих інформаційних систем. Київ, 2016. 218 с.

12. Терейковський І.А. Нейромережеві моделі, методи і засоби оцінювання параметрів безпеки інтернет-орієнтованих інформаційних систем : автореф. дис. ... док. техн. наук. Київ, 2015. 44 с.

13. William S. Network security essentials deep learning approach to network intrusion detection: applications and standards. 4th edition. New Jersey, USA. 2011. 417 p.

14. Комар М.П. Метод построения совокупного классификатора трафика информационно-телекоммуникационных сетей для иерархической классификации компьютерных атак. *Системи обробки інформації.* Харків, 2012. № 3. Том 1. С. 134–138.

15. Sriram S., Shashank A., Vinayakumar R., Soman K.P. DCNN-IDS : Deep Convolutional Neural Network based Intrusion Detection System. *TechRxiv, Preprint.* 2020. URL: https://doi.org/10.36227/techrxiv.12151734.v1.

16. Адамов О.С. Моделі і методи захисту кіберпростору на основі аналізу великих даних з використанням машинного навчання : автореф. дис. ... канд. техн. наук. Харків, 2019. 29 с.

17. Aloqailya M., Otoum S., Al Ridhawi I., Jararweh Y. An intrusion detection system for connected vehicles in smart cities. *Ad Hoc Networks*. 2019. Vol. 90. DOI: https://doi.org/10.1016/j.adhoc.2019.02.001.

18. Zhang C., Ruan F., Yin L., Chen X., Zhai L., Liu F. A Deep Learning Approach for Network Intrusion Detection Based on NSL-KDD Dataset. IEEE 13th International Conference on Anti-counterfeiting, Security, and Identification (ASID). 2019. P. 41–45. URL: https://doi.org/10.1109/ICASID.2019.8925239.

19. Mashayak S.A., Bombade B.R. Network Intrusion Detection Exploitation Machine Learning Strategies with the Utilization of Feature Elimination Mechanism. *International journal of computer sciences and engineering*. 2019. Vol. 7(5). P. 1292–1300. URL: https://doi.org/10.26438/ijcse/v7i5.12921300.

20. Almseidin M., Kovacs S. Intrusion Detection Mechanism Using Fuzzy Rule Interpolation. *Arxiv*. 2019. URL: https://arxiv.org/abs/1904.08790.

21. Mohammadi B., Sabokrou M. End-to-End Adversarial Learning for Intrusion Detection in Computer Networks. *Arxiv*. 2019. URL: https://arxiv.org/abs/1904.11577.

22. Suman C., Tripathy S., Saha S.. Building an Effective Intrusion Detection System using Unsupervised Feature Selection in Multi-objective Optimization Framework. *Arxiv*. 2019. URL: https://arxiv.org/abs/1905.06562.

23. Ferdowsi A., Saad W. Generative Adversarial Networks for Distributed Intrusion Detection in the Internet of Things. *Arxiv*. 2019. URL: https://arxiv.org/abs/1906.00567.

24. Hanselmann M., Strauss T., Dormann K., Ulmer H. CANet: An Unsupervised Intrusion Detection System for High Dimensional CAN Bus Data. *Arxiv.* 2019. URL: https://arxiv.org/abs/1906.02492.

25. Camacho J., García-Giménez J.M., Fuentes-García N.M., Maciá-Fernández G. Multivariate Big Data Analysis for Intrusion Detection: 5 steps from the haystack to the needle. *Arxiv.* 2019. URL: https://arxiv.org/abs/1906.11976.

26. Anton S.D.D., Sinha S., Schotten H.D. Anomaly-based Intrusion Detection in Industrial Data with SVM and Random Forests. *Arxiv*. 2019. URL: https://arxiv.org/abs/1907.10374.

27. Wu P., Guo H. LuNet: A Deep Neural Network for Network Intrusion Detection. *Arxiv*. 2019. URL: https://arxiv.org/abs/1909.10031.

28. Rawat S., Srinivasan A., Vinayakumar R. Intrusion detection systems using classical machine learning techniques versus integrated unsupervised feature learning and deep neural network. *Arxiv*. 2019. URL: https://arxiv.org/abs/1910.01114.

29. Gharib M., Mohammadi B., Dastgerdi S.H., Sabokrou M. AutoIDS: Auto-encoder Based Method for Intrusion Detection System. *Arxiv*. 2019. URL: https://arxiv.org/abs/1911.03306.

30. Zizzo G., Hankin C., Maffeis S., Jones K. Intrusion Detection for Industrial Control Systems: Evaluation Analysis and Adversarial Attacks. *Arxiv*. 2019. URL: https://arxiv.org/abs/1911.04278.

31. Zolanvari M., Teixeira M.A., Gupta L., Khan K.M., Jain R. Machine Learning-Based Network Vulnerability Analysis of Industrial Internet of Things. *IEEE Internet of Things Journal.* 2019. Vol. 6 (4). P. 6822–6834. URL: 10.1109/JIOT.2019.2912022.

32. Gwon H., Lee C., Keum R., Choi H. Network Intrusion Detection based on LSTM and Feature Embedding. *Arxiv.* 2019. URL: https://arxiv.org/abs/1911.11552.

33. Li G., Shen Y., Zhao P., Lu X., Liu J., Liu Y., Hoi S.C. H. Detecting cyberattacks in industrial control systems using online learning algorithms. *Neurocomputing.* 2019. Vol. 364. P. 338–348. URL: https://doi.org/10.1016/j.neucom.2019.07.031.

34. Soltani M., Siavoshani M.J., Jahangir A.H. A Content-Based Deep Intrusion Detection System. *Arxiv.* 2020. URL: https://arxiv.org/abs/2001.05009.

35. Abualkibash M. Machine Learning In Network Security Using Knime Analytics. *International Journal of Network Security & Its Applications (IJNSA).* 2019. Vol. 11 (5). URL: 10.5121/ijnsa.2019.11501.

36. Andalib A., Vakili V.T. An Autonomous Intrusion Detection System Using Ensemble of Advanced Learners. *Arxiv*. 2020. URL: https://arxiv.org/abs/2001.11936.

37. Alhajjar E., Maxwell P., Bastian N.D. Adversarial Machine Learning in Network Intrusion Detection Systems. *Arxiv*. 2020. URL: https://arxiv.org/abs/2004.11898.

38. Терейковська Л.О., Іванченко Є.В., Погорелов В.В. Метод адаптації глибокої нейронної мережі до розпізнавання комп'ютерних вірусів. *Комп'ютерно-інтегровані технології: освіта, наука виробництво.* Луцьк, 2019. № 35. С. 198–205.

39. Ludwig S.A. Applying a neural network ensemble to intrusion detection. *Journal of artificial intelligence and soft computing research.* 2019. Vol 9. P. 177–188. URL: https://doi.org/10.2478/jaiscr-2019-0002.

40. Geng X., Chen X., Zhu K.Q. MICK: A Meta-Learning Framework for Few-shot Relation Classification with Little Training Data. *Arxiv*. 2020. URL: https://arxiv.org/abs/2004.14164.

Babkin A.A., Kudin O.V. AN OVERVIEW OF NEURAL NET MODELS FOR INTRUSION DETECTION SYSTEM

The problem of information security is significant in the modern information society. Data protection on the Internet of Things and embedded systems has been the subject of recent publications. Personal and corporate data are potential vulnerabilities. Therefore, specialized software is needed to detect security threats and then automatically respond or inform users. The main purpose of such software is to identify threats in the real time network traffic. Intrusion detection systems are great example of such software. The aim is to detect potential attacks in computer networks. There are two main methods here: signature-based detection and anomaly-based detection. It is assumed that the behavior of an attacker on a computer network is different from the digital traces left by the regular user. Therefore, the problem of intrusion detection is usually reduced to the analysis of network traffic and the development of recognition methods. Anomaly-based approach requires some kind of machine learning methods for automatic anomaly detection and different types of neural networks are quite useful here. Features of such intrusion detection system are high learning speed and adaptability to new types of attacks. Typically, this data protection software works in real time and analyzes traffic both inner and external network requests. This paper discusses the models and methods of machine learning that are employed to solve the problem of automatic intrusion detection. We focus on works that offer methods for building hybrid systems based on a combination of neural networks with other machine learning methods. Such approaches have yielded significant results in the field of computer vision and machine translation. In addition, we outline a direction for further development of such models.

Key words: *Intrusion detection system, Neural Networks, Information Security, Hybrid System, Ensemble Learning.*