МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені І.І.МЕЧНИКОВА

(повне найменування вищого навчального закладу)

Факультет математики, фізики та інформаційних технологій

(повне найменування інституту, назва факультету (відділення))

Кафедра математичного забезпечення комп'ютерних систем

(повна назва кафедри (предметної, циклової комісії))

Дипломна робота

на здобуття ступеня вищої освіти «бакалавр»

(освітньо-кваліфікаційний рівень)

на тему «Система виявлення атак в комп'ютерних мережах на основі машинного навчання»

«Attack detection system in computer networks based on machine learning»

Виконала: студентка денної форми навчання

спеціальності 123 – Комп'ютерна інженерія .

(шифр і назва напряму підготовки, спеціальності)

Якушина Анастасія Олексіївна

(прізвище, ім'я, по-батькові)

Керівник к.ф.-м.н., доц. Шпінарева І. М

(науковий ступінь, вчене звання, прізвище та ініціали, **підпис**)

Рецензент к.техн.н., доц. Пенко В. Г.

(науковий ступінь, вчене звання, прізвище та ініціали)

Рекомендовано до захисту:

Протокол засідання кафедри

№ 11 від «11» 06 2021 р.

Завідувач кафедри

_____ Є.В. Малахов

(підпис) (прізвище, ініціали)

Захищено на засіданні ЕК № ___

протокол № __ від «___» ____ 2021 р.

Оцінка _____/_____/_____

(за національною шкалою, шкалою ECTS, бали)

Голова ЕК

_____ Н.Ф. Казакова

(підпис) (прізвище, ініціали)

Одеса - 2021

АНОТАЦІЯ

У дипломній роботі розроблено систему виявлення атак в комп’ютерних мережах на основі машинного навчання.

Кількість кібератак та порушень даних надзвичайно зросла у різних підприємствах, компаніях та галузях в результаті використання слабких сторін комп'ютерних систем. Парадигми систем виявлення вторгнень, які базуються на методах сигнатур, більше не є достатніми, оскільки не мають можливість виявляти нові атаки відсутні в базі знань, на відміну від методів виявлення аномальної поведінки мережевого трафіку в комп’ютерній мережі.

Метою роботи є створення системи виявлення атак в комп’ютерних мережах на основі машинного навчання.

В результаті проведених в роботі досліджень за методами виявлення атак в комп’ютерних мережах спроектована та реалізована система виявлення атак, що складається з модуля сенсору, модуля обробки пакетів, модуля аналізу та модуля сповіщення.

Модуль аналізу складається з двох складових: модуля виявлення атаки та модуля класифікації атаки. Запропонована модель використовує каскад з двох нейронних мереж: перша мережа перевіряє наявність атаки, а друга мережа класифікує атаку у разі її наявності. Для першої нейронної мережі використано гібридну нейронну мережі, яка складається з шарів згорткової нейронної мережі та рекурентної нейронної мережі з довгою короткочасною пам'яттю. Для другої – тільки згорткова нейрона мережа, оскільки вона показала більшу ефективність в задачі класифікації атаки.

Система виявлення атак навчена і протестована з використанням набору UNSW-NB15. Також система протестована в режимі реального часу за допомогою утиліти hping3 для моделювання атаки TCP SYN Flood. Всього відправлено 15000 пакетів, з яких система розпізнала 14182 як DoS-атаку, що склало 94,54% точності.

ABSTRACT

In this diploma work, an attack detection system in computer networks based on machine learning is developed.

Cyberattacks and data breaches have increased tremendously in various businesses, companies and industries as a result of exploitation of vulnerabilities in computer systems and networks. The signature-based intrusion detection paradigms are no longer sufficient since they do not have the ability to detect new attacks that are missed from the knowledge base, unlike anomalous behavior detection methods of network traffic in a computer network.

The aim of the work is to create an attack detection system in computer networks based on machine learning.

As a result of the research carried out in the work on methods of detection in computer networks, an attack detection system was designed and implemented, consisting of a sensor module, a packet processing module, an analysis module and a notification module.

The analysis module consists of two components: an attack detection module and an attack classification module. The proposed model uses a cascade of two neural networks: the first network checks for the presence of an attack, and the second network classifies the attack, if any. For the first neural network, a hybrid neural network was used, consisting of layers of a convolutional neural network and a recurrent neural network with long short-term memory. For the second one, only a convolutional neural network was used, since it has shown great efficiency in the problem of classifying an attack.

The intrusion detection system was trained and tested using the UNSW-NB15 dataset. The system was also tested in real time using the hping3 utility to simulate a TCP SYN Flood attack. A total of 15,000 packets were sent, of which the system recognized 14182 as a DoS attack, which was 94.54% of accuracy.

АННОТАЦИЯ

В дипломной работе разработана система обнаружения атак в компьютерных сетях на основе машинного обучения.

Количество кибератак и нарушений данных чрезвычайно возросла в различных предприятиях, компаниях в результате использования уязвимостей компьютерных систем и сетей. Парадигмы систем обнаружения вторжений, базирующиеся на методах сигнатур, больше не являются достаточными, поскольку не имеют возможность выявлять новые атаки, отсутствующие в базе знаний, в отличие от методов выявления аномального поведения сетевого трафика в компьютерной сети.

Целью работы является создание системы обнаружения атак в компьютерных сетях на основе машинного обучения.

В результате проведенных в работе исследований по методам обнаружения атак в компьютерных сетях спроектирована и реализована система обнаружения атак, состоящий из модуля сенсора, модуля обработки пакетов, модуля анализа и модуля оповещения.

Модуль анализа состоит из двух составляющих: модуля обнаружения атаки и модуля классификации атаки. Предложенная модель использует каскад из двух нейронных сетей: первая сеть проверяет наличие атаки, а вторая сеть классифицирует атаку в случае ее наличия. Для первой нейронной сети использована гибридная нейронная сеть, состоящей из слоев свёрточной нейронной сети и рекуррентной нейронной сети с длинной кратковременной памятью. Для второй – только свёрточная нейронная сеть, так как она показала большую эффективность в задачи классификации атаки.

Система обнаружения атак обучена и протестирована с использованием набора UNSW-NB15. Также система протестирована в режиме реального времени с помощью утилиты hping3 для моделирования атаки TCP SYN Flood. Всего отправлено 15000 пакетов, из которых система распознала 14182 как DoS-атаку, что составило 94,54% точности.

ЗМІСТ

ВСТУП

Зростаюча кількість мережевих пристроїв і сервісів підвищило важливість мережевої безпеки. Зростає потреба в захисних заходах, оскільки хакери запускають атаки, щоб паралізувати або вкрасти інформацію з комп'ютерних систем, підключених до мережі. Прикладами таких атак є передача шкідливих файлів і використання вразливостей систем. В таких атаках хакери взаємодіють з цільовою системою, викликаючи мережеву активність.

Кібератака є формою небезпечної атаки, яка швидко зростає як за кількістю зафіксованих атак, так і за ступенем їх шкоди для організацій та бізнесу. Дослідження [1, 2] класифікували методи кібератаки на дві основні категорії: пасивна атака та активна атака. Згідно з повідомленнями [3, 4], у 2019 році техніки кібератаки вважаються одними з найнебезпечніших методів атаки.

Мережеві системи виявлення вторгнень у мережі є важливими інструментами забезпечення мережевої інформаційної безпеки, а методи машинного навчання стають дедалі популярнішим рішенням для NIDS.

В даний час існує два основних методи виявлення кібератак: метод на основі сигнатур з використанням відомих знань та метод на основі аномалії, заснований на аналізі даних та статистиці для виявлення ненормальних поведінки у мережі [1, 2]. Основним недоліком методу на основі сигнатур, що є загальним для всіх підходів, що базуються на знаннях, є необхідність частого оновлення, щоб не відставати від потоку виявлених нових вразливостей.

Метод, заснований на аномалії, зазвичай базується на двох основних техніках, які включають машинне навчання та глибоке навчання для класифікації аномальної та нормальної поведінки. Однак із поступовою складністю мережевого середовища існуючі рішення, що використовують традиційні методи машинного навчання, не можуть повною мірою використовувати багату інформацію в даних мережевого трафіку через свою

єдину структуру. Що ще важливіше, це призводить до того, що існуючі NIDS володіють неповними знаннями про область виявлення вторгнень і унеможливлять досягнення високого рівня виявлення та хорошої стабільності в новому середовищі.

Метою роботи є створення системи виявлення атак в комп'ютерних мережах на основі машинного навчання.

Основними задачами даної дипломної роботи є:

- дослідження методів виявлення атак і їх характеристики;
- формування вимог до створюваної системи;
- вибір архітектури, технологій і засобів реалізації системи;
- реалізація алгоритму виявлення атак;
- реалізація системи виявлення атак;
- тестування обраного алгоритму на тестовій виборці;
- тестування системи в режимі реального часу.

ВИСНОВКИ

В ході виконання дипломної роботи спроектована, реалізована та протестована система, що виконує аналіз даних трафіку комп'ютерних мереж на предмет наявності атакуючого впливу. Для цього проведено огляд існуючих систем виявлення вторгнень і аналіз методів виявлення атак, та зроблено висновок про доцільність виконання даного проекту з метою вирішення основних недоліків цих систем.

Система складається з модуля сенсору, модуля обробки даних, модуля аналізу та модуля сповіщення. Для визначення атаки використовувався метод виявлення аномальної поведінки з використанням штучної нейронної мережі.

Модуль аналізу складається з двох складових: модуля виявлення атаки та модуля класифікації атаки. Запропонована модель використовує каскад з двох нейронних мереж: перша мережа перевіряє наявність атаки, а друга мережа класифікує атаку у разі її наявності. Для першої нейронної мережі, яка вирішує задачу виявлення аномалії, використано гібридну нейронну мережу, яка складається з шарів згорткової нейронної мережі та рекурентної нейронної мережі з довгою короткочасною пам'яттю. Для другої – тільки згорткова нейрона мережа, оскільки вона показала більшу ефективність в задачі класифікації атаки.

Для навчання використовувався набір UNSW-NB15, особлива риса якого є наявність пакетів мережевого трафіку, які належать до 9 різних класів атак. Критерієм вибору ефективної конфігурації були використані спеціалізовані метрики, які використовуються для тестування сучасних IDS систем.

Система виявлення атак протестована на тестовій виборці із UNSW-NB15, що містить більш ніж 50 тис. векторів. Також система протестована у режимі реального часу за допомогою утиліти hping3, яка використовувалася при моделюванні атаки TCP SYN Flood. Всього відправлено 15000 пакетів, з яких система розпізнала 14182 як DoS атаку, що складає 94,54% точності. З

метою не порушити роботу всієї мережі при виявленні атаки, система не блокує з’єднання, а сповіщає про атаку адміністратора, який сам вирішує що робити.

Подальша робота в цьому напрямку має вестися в декількох напрямках. По-перше, це використання розробленої системи виявлення атак всередині системи запобігання атак. По-друге, якісне доопрацювання – наприклад, можливість збереження результатів аналізу трафіку в окремий файл або вказану базу даних. Також можливе вдосконалення модулю сповіщення шляхом додавання різних методів сповіщення користувача про атаку. З точки зору інтерфейсу, можливо розробити графічну обертку над побудованою системи, в якої крім вже реалізованих функцій системи, додати також модуль статистики за певні проміжки часу та можливість фільтрації збереженої історії за вибраними ознаками.

За результатами даної роботи опубліковано тези на всеукраїнській конференції.

ПЕРЕЛІК ВИКОРИСТАННИХ ДЖЕРЕЛ

1. Markus R., et al., A survey of network-based intrusion detection data sets.Computers & security. – vol. 8, no. 6. – 2019. – pp. 147–167.
2. Ansam. Kh et al., Survey of intrusion detection systems: techniques, datasets and challenges,Cybersecurity. – vol. 20. – 2019. – pp. 2-20. [Електронний ресурс] – Режим доступу: https://doi.org/10.1186/s42400-019-0038-7.
3. Imperva, 2019 Cyberthreat Defense Report. – 2019. [Електронний ресурс] – Режим доступу: https://www.imperva.com/resources/reports/CyberEdge-2019-CDR-Reportv1.1.pdf.
4. Sophos, Sophos 2020 Threat Report. – 2020. [Електронний ресурс] – Режим доступу:https://www.sophos.com/en-us/medialibrary/PDFs/technical-papers/sophoslabs-uncut-2020-threat-report.pdf.
5. Pontarelli S. Traffic-aware design of a high-speed fpga network intrusion detection system. IEEE Transactions on Computers. – 62(11):2322–2334. – 2013.
6. SNORT / Snort team. San Jose : Cisco Systems Inc. – 2018. [Електронний ресурс] – Режим доступу: https://www.snort.org/.
7. Snort 2.1. Обнаружение вторжений : книга / Джей Бил [и др.]. М. : Бином-пресс. – Изд. 2. – 2006. – 656 с.
8. Гамаюнов Д.Ю. Современные некоммерческие средства обнаружения атак/ Гамаюнов Д.Ю., Смелянский Р. Л. // Программные системы и инструменты. Тематический сборник. – М.: Ф-т ВМиК МГУ. – 2002. – 20 с.
9. Suricata. Open Source IDS/IPs/NSM engine [Електронний ресурс] – Режим доступу: https://suricata-ids.org/.
10. Suricata как IPS // Хабр. – 2013. [Електронний ресурс] – Режим доступу: https://habr.com/post/192884/.
11. Bro: A system for detecting network intruders in real-time / Vern Paxson // Computer Networks. Amsterdam : Elsevier. – No. 31 (23-24). – 1999. – p. 2435
12. Краткий анализ решений в сфере СОВ и разработка нейросетевого детектора аномалий в сетях передачи данных// Хабр. – 2018. [Електронний

ресурс] – Режим доступу: https://habr.com/post/358200/.

13. OSSEC [Електронний ресурс] – Режим доступу: https://www.ossec.net/.
14. OSSEC and attacking through the firewall // Intrusion detection and firewall security. – 2016. [Електронний ресурс] – Режим доступу: https://www.cs.hioa.no/teaching/materials/MS004A/html/L65.en.pdf.
15. Debar H. An Introduction to Intrusion-Detection Systems. – 2009. [Електронний ресурс] – Режим доступу: https://www.researchgate.net/publication/228589845_An_Introduction_to_Intrusion-Detection_Systems.
16. Lunt T. A prototype real-time intrusion-detection expert system. In Proceedings of the 1988 Symposium on Security and Privacy. / Lunt T. and Jagannathan R. – 1988. – pp. 59–66.
17. Naji H. Asax: Software architecture and rule-based language for universal audit trail analysis. Proceedings of the Second ESORICS. / Naji H., Baudouin L.C., Aziz M., and Mathieu I. – 1992.
18. Garvey T. Model-based intrusion detection. In Proceedings of the 14th National Computer Security Conference. / Garvey T., Lunt T.– p. 372.
19. Helman P. Foundations of intrusion detection. In Proceedings of the Fifth Computer Security Foundations Workshop. / Helman P, Liepins G., and Richards W. – p. 114.
20. Helman P. Statistical foundations of audit trail analysis for the detection of computer misuse. IEEE Transactions on Software Engineering. / Helman P. Liepins G. – 19(9). – p. 886.
21. Javitz H. The SRI IDES statistical anomaly detector. In Proceedings of the IEEE Symposium on Research in Security and Privacy. / Javitz H., Valdes A. – p. 316.
22. Chalapathy R. Deep learning for anomaly detection: A survey. / Chalapathy R., Chawla S. – 2019.
23. Глубокое обучение / пер. с анг. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс. – 2018. – 652 с.
24. Прохоров В.Г. Использование сверточных нейронных сетей для распознавания рукописных символов // УкрПрог, Киев. – 2008.

25. LSTM — нейронная сеть с долгой краткосрочной памятью // Neurohive. – 2018. [Електронний ресурс] – Режим доступу: https://neurohive.io/ru/osnovy-data-science/lstm-nejronnaja-set/
26. Ojha, V.K. Metaheuristic Design of Feedforward Neural Networks: A Review of Two Decades of Research./ Ojha, V.K.; Abraham, A.; Snasel, V – 2017. – p. 60.
27. Sahu, S.K. A Detail Analysis on Intrusion Detection Datasets. In Proceedings of the 2014 IEEE International Advance Computing Conference (IACC). / Sahu, S.K.; Sarangi, S.; Jena, S.K. – 2014. – p. 1348.
28. S. Stolfo, KDD Cup 1999 Data. [Електронний ресурс] – Режим доступу: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
29. Tavallaee M. A detailed analysis of the KDD CUP 99 data set, IEEE Symposium on Computational Intelligence for Security and Defense Applications, CISDA./ Tavallaee M., Bagheri E., Lu W., Ghorbani A. A. – 2009. – pp. 1–6.
30. The UNSW-NB15 Dataset Description [Електронний ресурс] – Режим доступу: https://researchdata.edu.au/unsw-nb15-dataset/1425943
31. Moustafa N. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set), MilCIS 2015. / Moustafa N., Slay J. – 2015. – pp. 1–6.
32. Bernstein R. Systems and methods for detection of anomalous network behavior. / Bernstein R., Dulkin A – 2016.
33. Tensorflow [Електронний ресурс] – Режим доступу: https://www.tensorflow.org/.
34. Keras [Електронний ресурс] – Режим доступу: https://keras.io/.
35. Tshark [Електронний ресурс] – Режим доступу: https://www.wireshark.org/docs/man-pages/tshark.html.
36. Argus [Електронний ресурс] – Режим доступу: https://openargus.org/.
37. Monowar H. Network anomaly detection: methods, systems and tools. IEEE Communications Surveys & Tutorials. / Monowar H., Dhruba K., and Jugal K. – 2014. – pp. 303–336.
38. Gideon Creech and Jiankun Hu. A semantic approach to host-based intrusion

detection systems using contiguousand discontiguous system call patterns. IEEE Transactions on Computers. – 2014. – pp. 807–819.

39. Fairuz A. Evaluation of machine learning classifiers for mobile malware detection. Soft Computing. / Fairuz A., Feizollah A., Nor B., and Gani A. – 2016. – pp. 343–357.
40. Wang Y. Statistical Techniques for Network Security: Modern Statistically-Based Intrusion Detection and Protection: Modern StatisticallyBased Intrusion Detection and Protection. IGI Global. – 2008.
41. Якушина А.О., Шпінарева І.М. Виявлення аномалій в мережевому трафіку з використанням методів глубокого навчання / Інформатика, інформаційні системи та технології: тези доповідей вісімнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 23 квітня 2021 р. – Одеса, 2021. – с. 172-174.