

DOI: 10.18372/2310-5461.71.21454

УДК 004.056.5:004.89

Г. В. Пекуровський, канд. техн. наук

Національний університет «Київський авіаційний інститут»

orcid.org/0009-0006-1993-8615

e-mail: hlib.pekurovskyi@npp.kai.edu.ua

С. В. Лазаренко, д-р техн. наук, проф.

Національний університет «Київський авіаційний інститут»

orcid.org/0000-0003-3529-4806

e-mail: serhii.lazarenko@npp.kai.edu.ua

О. М. Кутінов,

Штаб 3 армійського корпусу

orcid.org/0009-0000-5559-502X

e-mail: alexeykutinov@i.ua

МЕТОД ВИЯВЛЕННЯ АНОМАЛЬНИХ HTTP-ЗАПИТІВ ЗАСОБАМИ РЕКУРЕНТНОЇ НЕЙРОННОЇ МЕРЕЖІ НА ОСНОВІ ІНДЕКСУ ПІДОЗРІЛОСТІ СЕСІЇ

Вступ

Протокол HTTP лежить в основі переважної більшості сучасних вебзастосунків, а кількість атак, спрямованих саме на цей рівень взаємодії (SQL-ін'єкція, міжсайтовий скриптинг, обхід каталогів, ін'єкція команд), продовжує зростати разом із поширенням електронної комерції та хмарних сервісів. Сигнатурні міжмережеві екрани вебзастосунків ефективно блокують відомі шаблони атак, проте погано узагальнюються на нові або злегка видозмінені варіанти шкідливих запитів і потребують постійного оновлення бази сигнатур, що обмежує їхню практичну ефективність проти раніше не бачених атак.

Рекурентні нейронні мережі природно підходять для аналізу тексту HTTP-запиту, оскільки він являє собою послідовність символів із локальними синтаксичними залежностями: структура шляху, розділові символи параметрів, лапки та спеціальні символи в ін'єкційних корисних навантаженнях. Разом з тим більшість наявних підходів обмежуються класифікацією окремого запиту, тоді як в операційній практиці рішення про реагування ухвалюється не щодо одного запиту, а щодо сеансу роботи користувача загалом. Поодинокий підозрілий запит серед десятків нормальних може бути хибним спрацюванням моделі, тоді як серія помірно підозрілих запитів у межах одного сеансу, типова поведінка автоматизованого сканера вразливостей, є значно небезпечнішою, навіть якщо жоден із запитів окремо не перевищує порогу впевненості класифікатора.

Метою роботи є розробка рекурентної нейронної мережі для виявлення аномальних HTTP-

запитів на рівні окремого запиту, а також метрики індексу підозрливості сесії (ІПС), що агрегує оцінки аномальності запитів у межах сеансу з урахуванням як поодинокого небезпечного запиту, так і розподіленої підозрілої активності, і забезпечує калібровану оцінку сеансу в цілому, придатну для прийняття операційних рішень.

Постановка проблеми

Більшість існуючих рішень на основі рекурентних нейронних мереж розв'язують задачу виявлення аномалій HTTP-трафіку як задачу класифікації окремого запиту: мережа отримує на вхід один запит і повертає оцінку його аномальності. Проте в операційній практиці систем захисту вебзастосунків рішення про реагування (блокування сесії, посилений моніторинг, сповіщення адміністратора) ухвалюється не щодо ізольованого запиту, а щодо сеансу роботи користувача в цілому. Це створює розрив між рівнем, на якому працює класифікатор, і рівнем, на якому приймається операційне рішення.

Цей розрив породжує дві протилежні проблеми, які класифікація окремого запиту не розрізняє. По-перше, поодинокий запит із високою оцінкою аномальності серед десятків нормальних запитів сеансу може бути хибним спрацюванням моделі, і автоматичне реагування на нього підвищує частку хибних тривог. По-друге, серія помірно підозрілих запитів у межах одного сеансу (типознака автоматизованого сканування вразливостей чи розподіленого перебору параметрів) залишається непоміченою, якщо жоден із запитів окремо не перевищує порогу впевненості класифікатора, хоча сукупна поведінка сеансу вже свідчить про зловмисну активність.

Розв'язання цієї проблеми потребує механізму агрегації посимвольних оцінок аномальності запитів у межах сеансу до єдиної каліброваної оцінки, придатної для операційного рішення, причому такий механізм має одночасно реагувати на найбільш небезпечний запит сеансу та на розподілену підозрілу активність, що не концентрується в жодному окремому запиті. Додатковим ускладненням є те, що загальновживаний датасет CSIC 2010, на якому традиційно перевіряються подібні підходи, не містить природної мультизапитової сесійної структури, бо записи датасету є незалежними запитами без ідентифікатора сеансу, що унеможливує безпосереднє навчання й перевірку агрегувальної метрики без додаткового етапу конструювання сесій із наявних запитів і міток.

Аналіз останніх досліджень і публікацій

Застосування рекурентних нейронних мереж для виявлення веб-атак на рівні окремого HTTP-запиту досліджено в роботі [1], де дві рекурентні мережі навчаються без учителя окремо на шляху та на параметрах запиту, після чого їхні виходи використовуються як ознаки для класифікатора; підхід перевірено на скороченій версії датасету CSIC 2010 з високою точністю виявлення.

Загальніша здатність рекурентних мереж моделювати послідовності мережевої взаємодії продемонстрована в роботі [2], де послідовності мережових потоків токенизуються як «речення», що дозволяє мовній моделі виявляти нетипові комунікації між вузлами мережі без прив'язки до конкретного протоколу прикладного рівня.

Важливість аналізу поведінки саме на рівні сеансу користувача, а не окремого запиту, підкреслено в роботі [3], де систему WebLearner побудовано на основі глибокого навчання для виявлення аномальної поведінки вебкористувачів у межах сеансів перегляду. Автори показують, що ознаки послідовності запитів сеансу суттєво підвищують якість виявлення порівняно з аналізом ізольованих запитів.

Інтерес до гібридних архітектур, що поєднують згорткові та рекурентні шари для аналізу структури URL і корисного навантаження HTTP-запиту, підтверджується й у новіших роботах. У [4] запропоновано архітектуру, що поєднує гілку ознак URL, згорткову гілку для аналізу корисного навантаження та двонаправлений рекурентний шар із механізмом самоуваги, апробовану на датасеті CSIC 2010.

Методологію генерації датасету HTTP DATASET CSIC 2010 описано в роботі [5]. Трафік згенеровано автоматично для вебзастосунку електронної комерції з використанням реальних довідкових значень параметрів, а аномальні запити

охоплюють статичні атаки (запити до прихованих чи неіснуючих ресурсів), динамічні атаки, що видозмінюють значення параметрів (SQL-ін'єкція, міжсайтовий скриптинг, переповнення буфера), та неумисно неправильні запити.

Водночас у роботі [6] зазначено, що датасет CSIC 2010 містить трафік лише до одного вебзастосунку з одного вузла і не відтворює мультизапитову структуру реального сеансу користувача, що обмежує його придатність для задач, де одиницею аналізу є саме сеанс, а не окремий запит. Це спостереження безпосередньо враховано в методології цієї роботи.

Постановка завдання

Для розв'язання окресленої проблеми в роботі необхідно розробити рекурентну нейронну мережу, яка класифікує окремий HTTP-запит як звичайний або аномальний на основі символьного представлення його тексту, і навчити її на реальному наборі даних HTTP DATASET CSIC 2010.

Оскільки датасет CSIC 2010 не містить природної мультизапитової сесійної структури, необхідно розробити методику конструювання сеансів із наявних запитів і міток датасету, що відтворює типові сценарії поведінки користувача (як легітимної, так і зловмисної) і на цій основі сформувати набір сеансів для подальшої перевірки агрегувальної метрики.

Далі необхідно запропонувати метрику агрегації посимвольних оцінок аномальності запитів у межах сеансу до єдиної каліброваної оцінки, яка одночасно враховує найбільш небезпечний запит сеансу та розподілену підозрілу активність і залишається в межах діапазону $[0, 1]$ незалежно від довжини сеансу. Вибір вагових коефіцієнтів запропонованої метрики має бути обґрунтований співвідношенням вартості помилок першого й другого роду, а стійкість результату перевірена аналізом чутливості до варіювання цих параметрів.

Насамкінець необхідно порівняти якість запропонованої метрики із базовими правилами агрегації сеансу на сконструйованому наборі сеансів, щоб визначити практичні переваги й обмеження запропонованого підходу порівняно з простішими альтернативами.

Виклад основного матеріалу дослідження

1. Методологія дослідження

У дослідженні використано реальний датасет HTTP DATASET CSIC 2010 [7], що містить 61065 HTTP-запитів до вебзастосунку електронної комерції, з яких 36000 позначено як звичайні та 25065 як аномальні. Для кожного запиту з полів методу, URL та тіла запиту (за наявності) сформо-

вано єдиний текстовий рядок, що відповідає рядку HTTP-запиту разом із параметрами чи вмістом POST-тіла; саме цей текст подається на вхід рекурентної нейронної мережі.

Датасет CSIC 2010 не містить природної мультизапитової сесійної структури, бо кожному запису відповідає унікальний ідентифікатор сеансу (JSESSIONID), тобто в первинних даних кожен запит є окремим ізольованим сеансом з одного запиту. Оскільки метрика індексу підозрілості сесії потребує сеансів із кількома запитами, реальні запити та реальні мітки датасету згруповано у сконструйовані сеанси за трьома сценаріями поведінки користувача:

1. Звичайна сесія: усі запити сеансу є звичайними і вибираються переважно в межах одного розділу вебзастосунку (публічна частина, особистий кабінет, статичні ресурси), що імітує послідовну навігацію користувача.

2. Сесія з поодиноким атакою: до послідовності звичайних запитів на випадковій позиції додається один реальний аномальний запит, що імітує поодинокі спроби експлуатації на фоні звичайної навігації.

3. Сесія сканування: від 30 % до 60 % запитів сеансу є реальними аномальними запитами, перемішаними зі звичайними, що імітує поведінку автоматизованого сканера вразливостей.

Сеанси формуються з ймовірностями 0,60 для звичайної сесії, 0,20 для сесії з поодиноким атакою та 0,20 для сесії сканування; довжина сеансу становить від 6 до 22 запитів. Усього сконструйовано 3000 сеансів, що охоплюють 41781 запит (24795 запитів у звичайних сеансах, 8532 у сеансах з поодиноким атакою, 8454 у сеансах сканування); загальна частка аномальних запитів після групування становить 10,61 %. Оскільки реальних запитів у датасеті (61065) менше за кількість запитів у сконструйованих сеансах, вибірка запитів для сеансів здійснюється з поверненням, бо той самий реальний запит теоретично може з'явитися в кількох різних сеансах.

Розбиття на навчальну, валідаційну та тестову вибірки виконується на рівні сеансів (а не окремих запитів), щоб уникнути потрапляння запитів одного сеансу одночасно у навчальну і тестову вибірки. Спочатку 30 % сеансів виокремлюється у тестову вибірку зі стратифікацією за міткою сеансу (звичайний або аномальний), після чого з решти 70 % сеансів 15 % виокремлюється у валідаційну вибірку. У підсумку навчальна вибірка містить 24885 запитів, валідаційна 4407 запитів, тестова 12489 запитів.

Кожен запит кодується посимвольно: використовується фіксований словник із 95 друкованих символів ASCII (код 0 зарезервовано під символ

доповнення), а довжина послідовності обмежується 200 символами, що охоплює 83 % запитів датасету без обрізання; довші запити обрізаються, коротші доповнюються нулями.

Архітектура рекурентної нейронної мережі складається з трьох компонентів: шару вкладення символів (кожен символ відображається у вектор розмірності 16), рекурентного шару з тангенціальною функцією активації прихованого стану [8]

$$h_t = \tanh(x_t W_{xh} + h_{t-1} W_{hh} + b_h),$$

де h_t є прихованим станом розмірності 32, $\tanh$ є функцією активації, що стискає результат у діапазоні $[-1; 1]$, x_t є вектором вкладення символу на кроці t , W_{xh} , W_{hh} і b_h є матрицями та вектором вагових параметрів, і вихідного повнозв'язного шару із сигмоїдною функцією активації, що дає ймовірність належності запиту до класу аномальних. Підсумкове представлення запиту обчислюється як усереднення прихованих станів по всіх реальних (недоповнених) символах запиту, а не як прихований стан на останньому кроці послідовності: це усуває розмивання сигналу великою кількістю кроків доповнення в кінці коротких запитів, що є типовим для символьного кодування запитів різної довжини.

Параметри мережі навчаються методом зворотного поширення помилки в часі з оптимізатором Adam (темп навчання 0,01, розмір пакета 128), а функцією втрат є бінарна крос-ентропія. Навчання виконується з ранньою зупинкою за втратою на валідаційній вибірці (параметр очікування становить 3 епохи), а максимальна кількість епох дорівнює 15.

2. Метрика індексу підозрілості сесії

Наявні підходи до виявлення веб-атак здебільшого приймають рішення щодо кожного HTTP-запиту окремо, тоді як операційне рішення про реагування (блокування сеансу, підвищена перевірка, сповіщення адміністратора) зазвичай ухвалюється щодо сеансу користувача в цілому. Пряме перенесення порогового рішення з рівня запиту на рівень сеансу неоднозначне: правило максимуму (сеанс вважається аномальним, якщо хоча б один запит перевищує поріг) чутливе до поодинокого хибнопозитивного спрацювання в довгому сеансі, тоді як правило середнього значення розмиває поодинокі небезпечні спроби серед великої кількості звичайних запитів сеансу.

Для сеансу s з оцінками аномальності запитів $p_1, \dots, p_n$, отриманими від рекурентної нейронної мережі, обчислюються три величини: максимальна оцінка аномальності серед запитів сеансу $p_{\max}(s)$, середня оцінка аномальності по сеансу $p_{\text{mean}}(s)$ та частка запитів сеансу, оцінка яких перевищує поріг τ , позначена $\rho(s)$. Базова оцінка сеан-

су поєднує чутливість до поодинокого небезпечного запиту з чутливістю до розподіленої аномальної активності:

$$B(s) = \beta \cdot p_{\max}(s) + (1 - \beta) \cdot p_{\text{mean}}(s).$$

Індекс підозрілості сесії визначається формулою

$$\text{ІПС}(s) = B(s) + \alpha \cdot \rho(s) \cdot (1 - B(s)),$$

параметри якої наведені у табл. 1. Другий доданок є підсиленням базової оцінки за рахунок частки підозрілих запитів сеансу, що масштабується залишковою відстанню до максимального значення $(1 - B(s))$; завдяки цьому ІПС(s) завжди залишається у діапазоні від 0 до 1 за умови, що α та β належать проміжку від 0 до 1.

Таблиця 1

Параметри метрики ІПС

Параметр	Значення	Опис
$p_{\max}(s)$	–	Максимальна оцінка аномальності серед запитів сеансу (вихід рекурентної нейронної мережі)
$p_{\text{mean}}(s)$	–	Середня оцінка аномальності запитів сеансу
$\rho(s)$	–	Частка запитів сеансу з оцінкою, що перевищує поріг τ
β	0,7	Вага максимальної оцінки відносно середньої в базовій оцінці $B(s)$
α	0,6	Коефіцієнт підсилення внеску частки підозрілих запитів $\rho(s)$
τ	0,5	Поріг віднесення окремого запиту до підозрілих при обчисленні $\rho(s)$

Вибір ваги $\beta = 0,7$ обґрунтовується асиметрією вартості помилок другого роду: пропуск поодинокого запиту з успішною експлуатацією вразливості (наприклад, дієвої SQL-ін'єкції) має суттєво вищу вартість, ніж пропуск розподіленої низькоінтенсивної активності (наприклад, розвідувального сканування), яку зазвичай можливо виявити повторно при подальшій взаємодії. Приймаючи орієнтовне співвідношення вартості таких помилок як 7 до 3, отримуємо $\beta = 7/(7+3) = 0,7$. Експериментальна перевірка на тестових сеансах підтверджує цей вибір: значення F1-міри при пороговому значенні ІПС 0,5 досягає максимуму серед перевірених значень β з проміжку від 0,5 до 0,9 саме при $\beta = 0,7$ (0,8264 проти 0,8000 при $\beta = 0,5$ і 0,8103 при $\beta = 0,9$), тоді як площа під ROC-кривою монотонно зростає з ростом β через високу роздільну здатність мережі на рівні окремого запиту.

Коефіцієнт $\alpha = 0,6$ обирається так, щоб сеанс, у якому кожен запит перевищує поріг τ (тобто $\rho(s) = 1$), незалежно від значення базової оцінки $B(s)$ отримував ІПС(s) не нижче 0,6, гарантуючи, що сеанс із суцільною підозрілою активністю не буде занижено оцінений через поодинокі невпевнені оцінки мережі. Аналіз чутливості за α у проміжку від 0,3 до 0,9 показує, що площа під ROC-кривою залишається практично незмінною (від 0,9337 до 0,9338), а F1-міра коливається у вузькому діапазоні від 0,8230 до 0,8270, що

підтверджує стійкість ранжування сеансів за ІПС до конкретного значення α у цьому проміжку.

3. Програмна реалізація

Експериментальну частину роботи реалізовано мовою Python у вигляді єдиного скрипта без застосування зовнішніх фреймворків глибокого навчання: рекурентна нейронна мережа, включно з прямим проходом та зворотним поширенням помилки в часі, реалізована безпосередньо на бібліотеці `numru`, що забезпечує повну прозорість обчислень і незалежність відтворюваності результатів від версій сторонніх фреймворків. Для роботи з табличними даними використано `pandas`, для розбиття вибірок і обчислення метрик якості класифікації, зокрема прецизійності, повноти, F1-міри, ROC-AUC та матриць помилок, використано бібліотеку `scikit-learn` [9]. Побудову графіків виконано засобами `matplotlib` і `seaborn`.

Відтворюваність результатів забезпечується фіксацією початкового значення генератора випадкових чисел (`SEED = 42`) як для стандартного модуля `random`, так і для глобального генератора `numru`, що використовується на кожному етапі: конструюванні сеансів із реальних запитів, розбитті на навчальну, валідаційну та тестову вибірки, ініціалізації параметрів мережі та перемішуванні пакетів під час навчання. Повторний запуск скрипта на тому самому файлі `csic_database.csv` дає числово ідентичні результати, таблиці та графіки, що перевірено порівнянням протоколів двох незалежних запусків.

4. Експериментальні результати

На рівні окремого запиту рекурентна нейронна мережа досягає на тестовій вибірці (12489 запитів) значення площі під ROC-кривою 0,9647 та втрати бінарної крос-ентропії 0,1357; навчання завершилося ранньою зупинкою на 10-й епосі з відновленням стану мережі, отриманого на 7-й епосі (втрата на валідаційній вибірці 0,1377), що свідчить про високу роздільну здатність символного представлення запитів CSIC 2010 для рекурентної архітектури навіть без застосування складніших рекурентних елементів.

Тестова вибірка охоплює 900 сеансів. У середньому для сеансів кожного сценарію отримано такі значення максимальної оцінки, середньої оцінки, частки підозрілих запитів та ІПС відповідно: для звичайних сеансів 0,163/0,048/0,009/0,131;

для сеансів з поодинокією атакою 0,598/0,087/0,048/0,456; для сеансів сканування 0,968/0,297/0,263/0,801. Отримані значення узгоджуються з очікуваною поведінкою метрики: сеанси з поодинокією атакою мають помітно нижчу середню оцінку і частку підозрілих запитів порівняно з сеансами сканування, тоді як максимальна оцінка зростає в обох типах аномальних сеансів відносно звичайних.

Для порівняння ІПС із базовими правилами агрегації пороги правила максимуму та правила середнього підібрано на валідаційних сеансах за критерієм максимізації F1-міри (0,663 та 0,049 відповідно), тоді як для ІПС використано його змістовний серединний поріг 0,5, оскільки метрику спеціально побудовано каліброваною в діапазоні від 0 до 1 незалежно від складу сеансу. Результати наведено у табл. 2.

Таблиця 2

Результати виявлення аномальних сеансів на тестовій вибірці

Правило	Прецизійність	Повнота	F1-міра	ROC-AUC	FPR
Правило максимуму	0,9704	0,7218	0,8278	0,9363	0,0149
Правило середнього	0,6435	0,8402	0,7288	0,8706	0,3147
ІПС	0,9340	0,7410	0,8264	0,9338	0,0354

Правило середнього значення після підбору порогу забезпечує найвищу повноту (0,8402), проте ціною істотно вищої частки хибнопозитивних спрацювань (0,3147) і найнижчого значення F1-міри серед трьох правил, що робить його малопридатним для практичного застосування: майже третина звичайних сеансів була б помилково позначена як аномальна. Правило максимуму після підбору порогу на валідаційних сеансах демонструє найвищу прецизійність (0,9704) і найнижчу частку хибнопозитивних спрацювань (0,0149), а також дещо вищу F1-міру, ніж запропонований ІПС (0,8278 проти 0,8264).

Разом з тим цей результат правила максимуму досягається лише за умови ретельного підбору порогу 0,663 на окремій валідаційній вибірці розмічених сеансів; за відсутності такої вибірки природний поріг 0,5 для правила максимуму дає істотно нижчу прецизійність. Запропонований ІПС натомість використовує змістовний, наперед визначений поріг 0,5 і за цим порогом демонструє прецизійність 0,9340, повноту 0,7410 та F1-міру 0,8264, практично не поступаючись найкращому з підібраних базових правил, водночас істотно перевершуючи правило середнього значення за F1-мірою (0,8264 проти 0,7288) та частку хибно-

позитивних спрацювань (0,0354 проти 0,3147). Отже, головна практична перевага ІПС полягає не значення суттєво нижча вже за повноти понад 0,4, що узгоджується з висновком про розмивання поодинокієї атаки при усередненні оцінок по довгому сеансу. в перевищенні найкращого підбраного базового правила, а в тому, що зіставний рівень якості досягається без потреби в емпіричному підборі порогу для кожного конкретного набору даних.

ROC-криві (рис. 1) підтверджують близькість роздільної здатності правила максимуму та ІПС (площа під кривою 0,9363 і 0,9338 відповідно) і помітне відставання правила середнього значення (0,8706), особливо в області низьких значень частки хибнопозитивних спрацювань, що є найбільш релевантною для практичного застосування областю.

Криві прецизійності та повноти (рис. 2) демонструють практично однакову форму для правила максимуму та ІПС на всьому діапазоні повноти, тоді як крива правила середнього значення суттєво нижча вже за повноти понад 0,4, що узгоджується з висновком про розмивання поодинокієї атаки при усередненні оцінок по довгому сеансу.

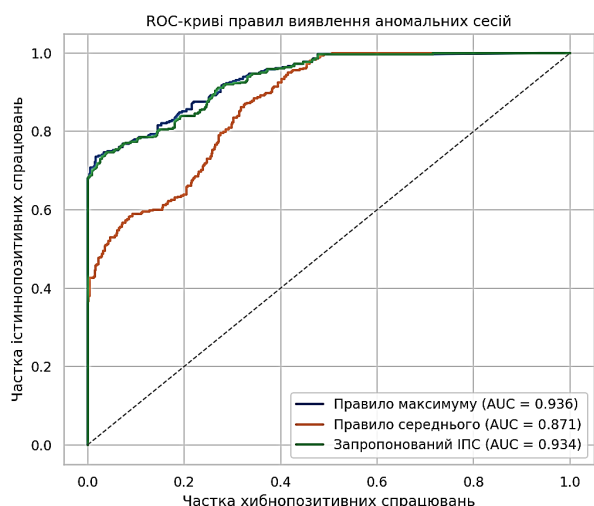


Рис. 1. ROC-криві правил виявлення аномальних сесій

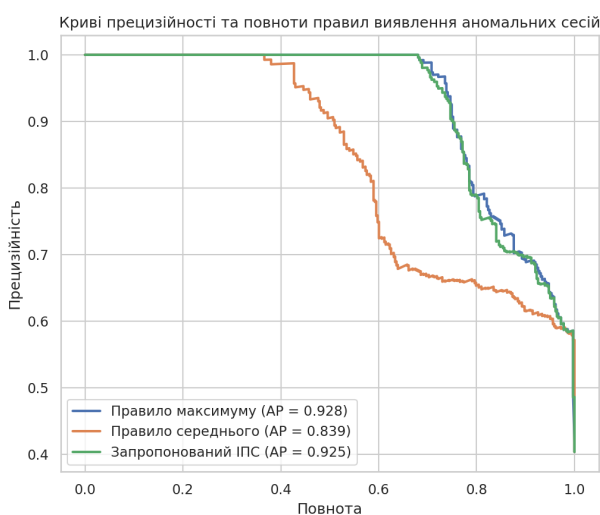


Рис. 2. Криві прецизійності та повноти правил виявлення аномальних сесій

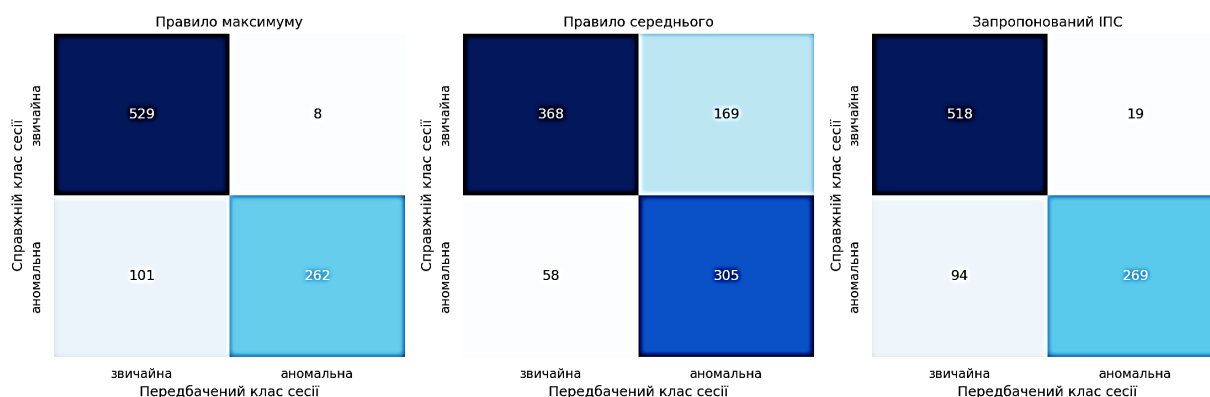


Рис. 3. Матриці помилок правил виявлення аномальних сесій за підібраними порогами

Матриці помилок (рис. 3) деталізують структуру помилок кожного правила: правило середнього значення допускає 169 хибнопозитивних спрацювань на 537 звичайних сеансах, тоді як ІПС і правило максимуму допускають лише 19 і 8 хибнопозитивних спрацювань відповідно; водночас правило максимуму пропускає 101 аномальний сеанс, а ІПС 94, тобто дещо менше.

Зведена діаграма (рис. 4) узагальнює всі п'ять показників з табл. 2 в одному вигляді і наочно демонструє головний компроміс між правилами: правило середнього значення виграв лише за повнотою, але суттєво програє за прецизійністю та особливо за часткою хибнопозитивних спрацювань (FPR), де його стовпчик помітно вищий за інші два правила. Правило максимуму та запропонований ІПС натомість дають майже однакові за висотою стовпчики для прецизійності, F1-міри та ROC-AUC, і обидва тримають FPR на низькому рівні, що ще раз підтверджує висновок: ІПС не поступається найкращому підібраному базовому правилу за жодним із показників, водночас явно перевершуючи наївне усереднення оцінок по сесії.

Рис. 5 ілюструє формулу ІПС для типових значень базової оцінки $B(s)$, усереднених за кожним сценарієм сеансу, як функцію частки підозрілих запитів $p(s)$; вертикальні пунктирні лінії позначають середнє спостережене значення $p(s)$ для кожного сценарію на тестовій вибірці. Графік демонструє, що зростання частки підозрілих запитів підсилює ІПС тим сильніше, чим нижча базова оцінка сеансу, забезпечуючи додаткове зростання оцінки саме для сеансів з розподіленою підозрілою активністю, не даючи знизити її при помірних значеннях $p_{\max}$ та p_{mean} .

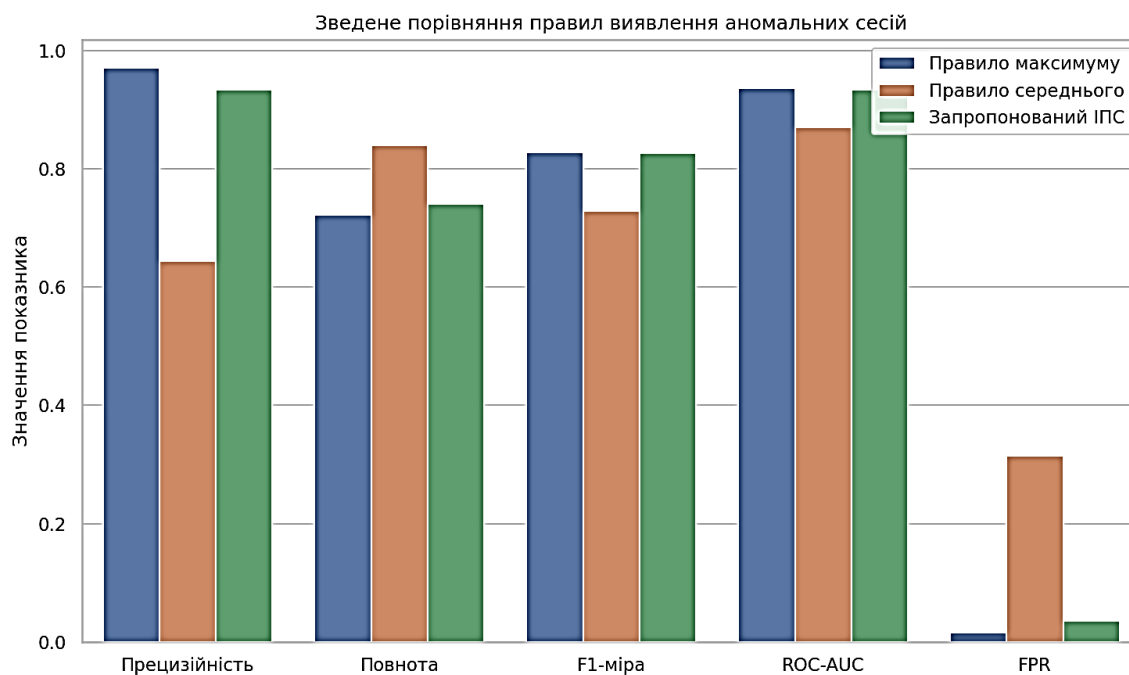


Рис. 4. Зведене порівняння правил виявлення аномальних сеансів за п'ятьма показниками

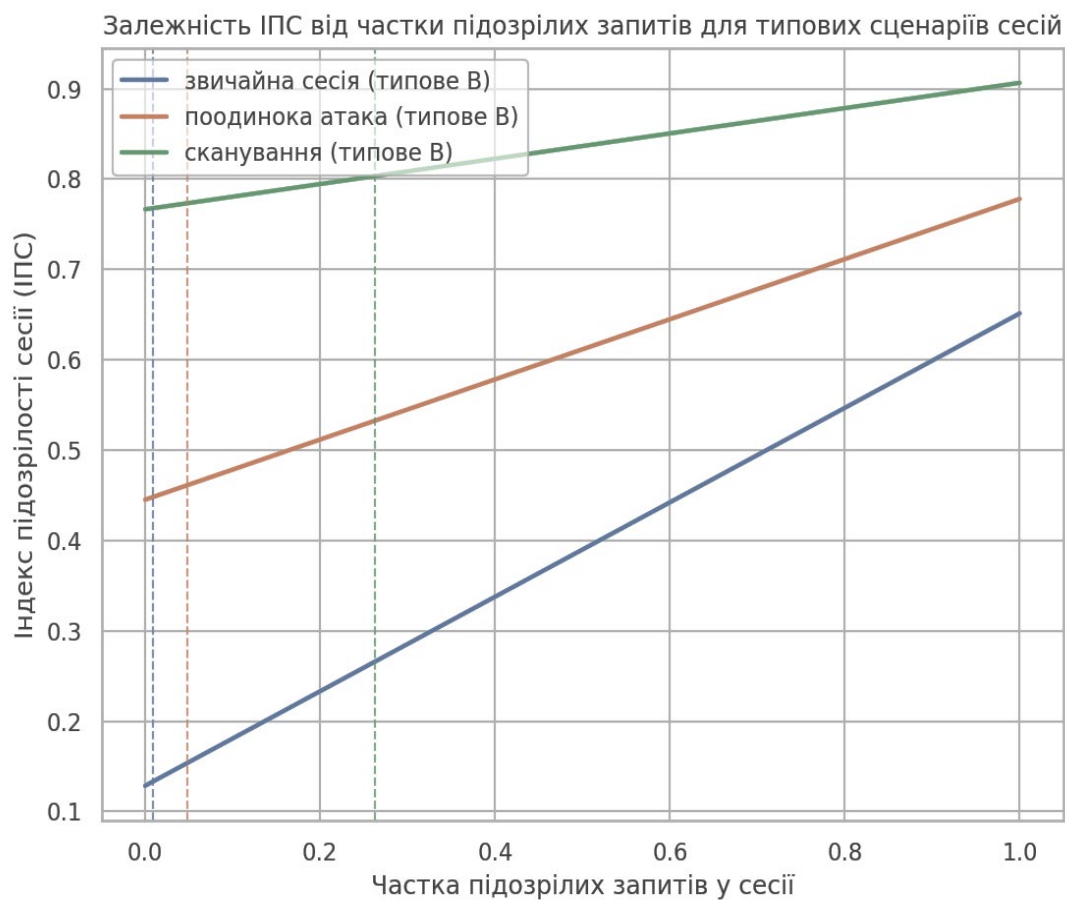


Рис. 5. Залежність ІПС від частки підозрілих запитів у сесії для типових значень базової оцінки трьох сценаріїв сеансів

5. Обговорення результатів

Отримані результати варто інтерпретувати з урахуванням ключового методологічного обмеження роботи: датасет CSIC 2010 не містить природної мультизапитової сесійної структури, тому сеанси в цьому дослідженні сконструйовано з реальних запитів і реальних міток датасету за трьома заданими сценаріями поведінки, а не запозичено з журналів реальних багатапитових сеансів користувачів. Це означає, що розподіл довжини сеансів, частка сеансів кожного сценарію та розташування аномальних запитів у сеансі є дослідницьким рішенням, а не спостереженою характеристикою реального трафіку, а тому кількісні значення показників можуть відрізнятися на реальних журналах вебсервера з іншими пропорціями сценаріїв поведінки користувачів.

Друге обмеження пов'язане зі способом формування сеансів із скінченного пулу реальних запитів (61065 записів) з поверненням, унаслідок чого той самий реальний запит теоретично міг увійти до кількох різних сконструйованих сеансів. Хоча кожен окремий запит і кожна мітка при цьому залишаються повністю реальними, це створює певну статистичну залежність між сеансами, що варто врахувати при перенесенні кількісних результатів на інші умови.

Третє спостереження стосується самого порівняння з базовими правилами: на датасеті CSIC 2010 окремі аномальні запити виявляються рекурентною мережею з високою впевненістю (площа під ROC-кривою на рівні запиту 0,9647), унаслідок чого навіть просте правило максимуму після підбору порогу на валідаційній вибірці демонструє результат, зрівняний із запропонованим ІПС. Це не применшує практичної цінності ІПС, оскільки, як показано вище, ця цінність полягає передусім у відсутності потреби в емпіричному підборі порогу для кожного конкретного набору даних і в кращій керованості частки хибнопозитивних спрацювань порівняно з правилом середнього значення, а не в беззастережному перевищенні найкращого підібраного базового правила за всіма показниками одночасно. Слід очікувати, що перевага ІПС над правилом максимуму стане більш вираженою на даних, де окремі запити атаки розпізнаються моделлю менш впевнено, наприклад через більшу подібність корисного навантаження до звичайного трафіку.

Перспективними напрямками подальших досліджень є перевірка запропонованої метрики на журналах реального вебсервера з природною мультизапитовою сесійною структурою, дослідження адаптивного підбору ваг β та α відповідно до конкретної моделі вартості помилок організації, а також розширення рекурентної архітек-

тури додатковими шарами чи двонапрямленою обробкою послідовності для запитів, довжина яких перевищує обраний поріг обрізання в 200 символів.

Висновки

У роботі розроблено рекурентну нейронну мережу для виявлення аномальних HTTP-запитів на основі символічного представлення тексту запиту, реалізовану безпосередньо засобами *numpy* без застосування зовнішніх фреймворків глибокого навчання. На реальному датасеті HTTP DATASET CSIC 2010 мережа досягає площі під ROC-кривою 0,9647 на рівні окремого запиту.

Запропоновано метрику індексу підозрілості сесії (ІПС), що агрегує оцінки аномальності запитів рекурентної мережі в межах сеансу користувача, поєднуючи чутливість до поодинокого небезпечного запиту з чутливістю до розподіленої підозрілої активності, і залишається каліброваною в діапазоні від 0 до 1 незалежно від складу сеансу. Обґрунтовано вибір параметрів метрики ($\beta = 0,7$, $\alpha = 0,6$) на основі асиметрії вартості помилок і підтверджено стійкість ранжування сеансів до варіювання цих параметрів у широкому діапазоні значень.

Порівняння на 900 тестових сеансах, сконструйованих із реальних запитів і міток CSIC 2010, показало, що ІПС за фіксованим порогом 0,5 забезпечує F1-міру 0,8264 та частку хибнопозитивних спрацювань 0,0354, істотно перевершуючи правило середнього значення (F1-міра 0,7288, частка хибнопозитивних спрацювань 0,3147) і практично не поступаючись правилу максимуму, підібраному за окремою валідаційною вибіркою (F1-міра 0,8278). Головна практична цінність запропонованої метрики полягає в тому, що зрівняний із найкращим підібраним базовим правилом рівень якості досягається без потреби в емпіричному підборі порогу для кожного конкретного набору даних, що є суттєвою перевагою в умовах обмеженої кількості розмічених сеансів.

Основним обмеженням роботи є конструювання сеансів із реальних, але позбавлених природної сесійної структури запитів датасету CSIC 2010, що визначає напрям подальших досліджень: перевірку метрики на журналах реального вебсервера з природними багатапитовими сеансами користувачів.

ЛІТЕРАТУРА

- [1] Liang J., Zhao W., Ye W. Anomaly-Based Web Attack Detection: A Deep Learning Approach // Proceedings of the 2017 VI International Conference on Network, Communication and Computing (ICNCC 2017). New York : ACM, 2017. P. 80–85.

- [2] Radford B. J., Apolonio L. M., Trias A. J., Simpson J. A. Network Traffic Anomaly Detection Using Recurrent Neural Networks // arXiv preprint. 2018. arXiv:1803.10769.
- [3] Gui J., Chen Z., Yu X., Lumezanu C., Chen H. Anomaly Detection on Web-User Behaviors Through Deep Learning // Security and Privacy in Communication Networks : SecureComm 2020 : Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 335. Cham : Springer, 2020.
- [4] Luo K., Chen Y. A Novel Hybrid Neural Network Approach Incorporating Convolution and LSTM With a Self-Attention Mechanism for Web Attack Detection // Applied Intelligence. 2024.
- [5] Torrano-Gimenez C., Perez-Villegas A., Alvarez Maranon G. An Anomaly-Based Approach for Intrusion Detection in Web Traffic // Journal of Information Assurance and Security. 2010. Vol. 5, No. 4. P. 446–454.
- [6] Pevný T., Dedič M. Nested Multiple Instance Learning in Modelling of HTTP Network Traffic // arXiv preprint. 2020. arXiv:2002.04059.
- [7] Instituto de Tecnologías Físicas y de la Información (ITEFI), CSIC. HTTP DATASET CSIC 2010. Madrid, 2010. URL: <https://www.tic.itefi.csic.es/dataset/>
- [8] Elman J. L. Finding Structure in Time // Cognitive Science. 1990. Vol. 14, No. 2. P. 179–211.
- [9] Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay É. Scikit-learn: Machine Learning in Python // Journal of Machine Learning Research. 2011. Vol. 12. P. 2825–2830.

Пекуровський Г. В., Лазаренко С. В., Кутінов О. М.

МЕТОД ВИЯВЛЕННЯ АНОМАЛЬНИХ HTTP-ЗАПИТІВ ЗАСОБАМИ РЕКУРЕНТНОЇ НЕЙРОННОЇ МЕРЕЖІ НА ОСНОВІ ІНДЕКСУ ПІДОЗРІЛОСТІ СЕСІЇ

У роботі досліджується задача виявлення аномальних HTTP-запитів у вебтрафіку засобами рекурентної нейронної мережі, що є актуальним напрямом захисту вебзастосунків в умовах постійного зростання кількості й різноманітності атак на прикладному рівні. Мережа навчається на реальному наборі даних HTTP DATASET CSIC 2010 і класифікує окремі запити як звичайні або аномальні на основі символічного представлення тексту запиту, що дозволяє виявляти як відомі, так і раніше не бачені варіанти шкідливих корисних навантажень без залежності від сигнатурної бази. Оскільки операційні рішення про реагування на вебзастосунковому рівні ухвалюються здебільшого не щодо окремого запиту, а щодо сеансу роботи користувача в цілому, у роботі запропоновано метрику індексу підозрілості сесії (ІПС), що агрегує посимвольні оцінки аномальності запитів сеансу з урахуванням як найбільш небезпечного окремого запиту, так і розподіленої підозрілої активності, що не концентрується в жодному конкретному запиті. Метрика забезпечує калібровану оцінку сеансу в діапазоні від 0 до 1 незалежно від його довжини, а вагові коефіцієнти обґрунтовано співвідношенням вартості помилок першого й другого роду. Оскільки датасет CSIC 2010 не містить природної мультизапитової сесійної структури, сесії сконструйовано з реальних запитів і реальних міток датасету за трьома сценаріями поведінки користувача, що охоплюють як легітимну активність, так і типову поведінку автоматизованого сканера вразливостей. Реалізацію рекурентної мережі виконано без використання сторонніх фреймворків глибокого навчання, що забезпечує повний контроль над обчислювальним процесом і прозорість архітектури. Результати застосування ІПС порівнюються з базовими правилами оцінювання сеансу за максимальним і середнім значенням оцінок аномальності запитів, а стійкість результатів перевірено аналізом чутливості до варіювання параметрів метрики.

Ключові слова: виявлення аномалій; HTTP-запити; нейронна мережа; безпека вебзастосунків; індекс підозрілості сесії; CSIC 2010; дисбаланс класів.

Pekurovskyi H., Lazarenko S., Kutinov O.

METHOD FOR DETECTING ANOMALOUS HTTP REQUESTS USING A RECURRENT NEURAL NETWORK BASED ON A SESSION SUSPICION INDEX

This paper investigates the task of detecting anomalous HTTP requests in web traffic using a recurrent neural network, which remains a relevant direction in web application security given the steadily growing volume and diversity of application-layer attacks. The network is trained on the real HTTP DATASET CSIC 2010 and classifies individual requests as normal or anomalous based on the character-level representation of the request text, enabling detection of both known and previously unseen malicious payload variants without dependence on a signature database. Since operational response decisions at the web application level are typically made not with respect to an individual request

but with respect to the user session as a whole, this paper proposes a session suspicion index (SSI) metric that aggregates character-level anomaly scores of a session's requests, accounting for both the single most dangerous request and distributed suspicious activity that is not concentrated in any individual request. The metric provides a calibrated session-level score within the range from 0 to 1 regardless of session length, and its weighting coefficients are justified based on the cost ratio of type I and type II errors. Since the CSIC 2010 dataset does not contain a natural multi-request session structure, sessions were constructed from the dataset's real requests and real labels according to three user behavior scenarios, covering both legitimate activity and typical automated vulnerability scanner behavior. The recurrent network was implemented without relying on third-party deep learning frameworks, which ensures full control over the computational process and transparency of the architecture. The SSI results are compared against baseline session scoring rules based on the maximum and mean values of request anomaly scores, and the robustness of the results is verified through a sensitivity analysis of the metric parameters.

Keywords: anomaly detection; HTTP requests; neural network; web application security; session suspicion index; CSIC 2010; class imbalance.

Дата першого надходження: 07.07.2026 р.

Дата прийняття до друку: 13.08.2026 р.

Дата публікації: 10.09.2026 р.