**УДК 004.8+056**

**Л. М. Куперштейн, к.т.н., доц., О.П. Войтович, к.т.н., доц.**
**А.Г. Буда, к.т.н., доц., О. С. Печенюк, маг.**

# ЗАХИСТУ ВЕБ-ДОДАТКУ ВІД XSS-АТАК НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

Із розвитком інформаційно-комунікаційних систем все частіше постає питання їх захисту адже з кожним роком виникають нові методи та нові типи атак, які можуть нанести шкоду. При створенні веб-додатків все більше розробників використовують мову програмування JavaScript. Щоправда її розповсюдження викликає проблему вразливості веб-сайту до різноманітних атак. Серед них однією з найбільш небезпечних та поширених є атака типу Cross-site scripting або XSS. У 2019 році атака даного типу була визнана найпоширенішою у світі, адже була використана у 39% кібератак на веб-додатки [1]. При цьому активно використовуються ряд підходів до виявлення та запобіганню XSS-атак. Серед них можна виділити фільтрування різних html-тегів та їх атрибутів у HTTP-запитах, а також використання спеціальних HTTP-заголовків [2]. Тому **актуальним** на сьогодні розробка нових інтелектуальних підходів до забезпечення захисту web-додатків від атак міжсайтового скриптингу.

Однією із найбільш перспективних технологій превентивного захисту від XSS атак є штучні нейронні мережі, яким притаманні потужні апроксимаційні здібності [3].

**Постановка задачі.** Для виявлення і попередження XSS-атак необхідно проаналізувати HTTP-заголовок, що надходить на сервер для ідентифікації заборонених символів, слів та словосполучень. Якщо система аналізу приймає рішення про відсутність шкідливого коду у запиті, то такий запит позначається як «негативний» і обробляється, в іншому випадку, позначається як «позитивний» та блокується, надсилається повідомлення про помилку. Оскільки заголовки є текстовими даними, то найбільш доцільним для їх аналізу буде застосувати рекурентні мережі.

Для **розв'язання задачі** було використано архітектуру seq2seq, яка складається з двох LSTM нейронних мереж. Для навчання нейронної мережі використано датасет із 21991 HTTP-запитів без аномалій та 1097 аномальних запитів загальним розміром 400 Мб. Вибірку поділено на навчальну та тестувальну у співвідношенні 7:3. Символьні дані HTTP-запитів для нейромережевої обробки подаються у вигляді ASCII-кодів. Результат моделювання показали високу точність класифікації (табл. 1). При цьому процес навчання без використання GPU тривав близько двох годин.

Таблиця 1. Значення основних метрик оцінювання нейромережі

| Метрика | Значення | Формула |
|---|---|---|
| Чутливість (Sensitivity/Recall/True Positive Rate) | 1 | TPR = TP / (TP + FN) |
| Специфічність (Specificity/True Negative Rate) | 0.9964 | SPC = TN / (FP + TN) |
| Точність (Accuracy) | 0.9976 | ACC = (TP + TN) / (TP + TN + FP + FN) |
| Показник позитивного прогнозування (Precision/Positive Predictive Value) | 0.9928 | PPV = TP / (TP + FP) |
| False Positive Rate | 0.0036 | FPR = FP / (FP + TN) |
| False Negative Rate | 0 | FNR = FN / (FN + TP) |
| F1 Score | 0.9964 | F1 = 2TP / (2TP + FP + FN) |
| Matthews Correlation Coefficient | 0.9946 | TP*TN - FP*FN / sqrt((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)) |

**Висновки**. Для розробки програмного засобу використано бібліотеку TensorFlow та мову програмування Python. Засіб інтегровано у веб-додаток та протестовано у лабораторних умовах на наперед запрограмованих XSS-вразливостях. На проведених тестах точність виявлення склала 100%.

**Література**

1. XSS the most widely-used attack method of 2019. URL: https://www.cloudpro.co.uk/it-infrastructure/security/ 8352/xss-the-most-widely-used-attack-method-of-2019

2. Куперштейн Л.М., Войтович О.П. та ін. Аналіз методів захисту від XSS атак // "Інформаційні технології та комп'ютерне моделювання"; матеріали статей Міжнародної науково-практичної конференції, м. Івано-Франківськ, 18-22 травня 2020 року. – Івано-Франківськ: Голіней О.М., С. 146-148.

3. Васюра А.С., Мартинюк Т.Б., Куперштейн Л.М. Методи та засоби нейроподібної обробки даних для систем керування. Монографія. − Вінниця: УНІВЕРСУМ−Вінниця, 2008. − 175 с.