# ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ МОДЕЛЕЙ ТРАНСФОРМЕРІВ ПРИ ІДЕНТИФІКАЦІЇ МЕРЕЖЕВИХ АТАК

*ас. В.О. Полторацький, д-р техн. наук, проф. С.Ю. Гавриленко, Національний технічний університет "Харківський політехнічний інститут", м. Харків*

У задачах класифікації мережевих атак моделі трансформери демонструють високу ефективність завдяки здатності знаходити складні залежності в даних [1]. Багатозадачність трансформерних моделей дозволяє їм, наприклад, одночасно прогнозувати наступну мережеву подію та визначати її аномальність. Однак такі моделі схильні до перенавчання через високу розмірність вхідних ознак та складність архітектури.

Класичним підходом для боротьби з перенавчанням є використання методу регуляризації штучних нейронних мереж. Одним із найбільш популярних методів є метод Dropout [2]. Але існує складність з визначенням оптимального коефіцієнта регуляризації Dropout.

Для підвищення ефективності роботи моделей досліджено можливість використання методу SpatialDropout, який знижує ризик перенавчання, краще зберігає просторові чи семантичні зв'язки між ознаками, що критично для обробки числових і категоріальних даних у задачах мережевої безпеки. Розроблено програмні моделі трансформерів з використанням двох методів регуляризації Dropout та SpatialDropout1D. Навчання проводилося з використанням графічного процесору GPU T4. У якості вихідних даних використано датасет про мережеві вторгнення UNSW- NB15, із 4 класами (Normal, Exploits, Fuzzers, Generic). Використано такі налаштування моделі: 2 блоки трансформера, batch size 128, до 30 епох із EarlyStopping.

Застосування SpatialDropout1D дозволило зменшити час навчання моделі на 11% завдяки використанню більш ефективних масок і стабільнішого градієнту, покращити показник loss-функції на 6%, збільшити точність класифікації на 1%.

**Список літератури: 1.** Gavrylenko S., Poltoratskyi V., Nechyporenko A., Intrusion detection model based on improved transformer. *Advanced Information Systems*, 2024, 8(1), p. 94–99. **2.** Salehin, I., Kang, D., Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain, *Electronics*, 2023, 12, p. 3182-3106.