DOI: 10.34185/1991-7848.itmm.2026.01.090

ПРОБЛЕМИ ОПТИМІЗАЦІЇ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ ДЛЯ БЕЗПЕРЕРВНОГО ЗАХИСТУ КОРПОРАТИВНИХ ІНФОРМАЦІЙНИХ СИСТЕМ

Рябенко В.І. [ORCID]

Дніпровський державний технічний університет, аспірант, Україна

Анотація. *У роботі досліджуються математичні та алгоритмічні проблеми оптимізації моделей машинного навчання (МН), що застосовуються для безперервного моніторингу та захисту сучасних корпоративних інформаційних систем (КІС). Зростання обсягів мережевого трафіку та перехід до архітектури нульової довіри (Zero Trust) вимагають від систем безпеки аналізу даних у режимі реального часу, що призводить до критичного збільшення обчислювального навантаження. Розглянуто ключові бар'єри впровадження МН, зокрема високу розмірність ознакового простору, проблему дисбалансу класів (аномалій порівняно з нормальним трафіком) та затримки під час інференсу. Обґрунтовано необхідність застосування методів зменшення розмірності та ансамблевих підходів для підвищення точності виявлення багатовекторних кібератак при одночасному зниженні частки хибних спрацьовувань. Запропоновано напрями математичного моделювання для мінімізації функції втрат в умовах динамічних змін у КІС.*

Ключові слова: *корпоративні інформаційні системи, машинне навчання, оптимізація алгоритмів, безперервний захист, архітектура нульової довіри, обчислювальна складність.*

Вступ. Еволюція корпоративних інформаційних систем, їх перехід до гібридних та хмарних середовищ вимагають відмови від традиційного захисту периметра на користь концепції безперервної верифікації (Continuous Adaptive Risk and Trust Assessment) [1]. У цьому контексті інструменти машинного навчання стають безальтернативним базисом для виявлення складних аномалій і багатоступеневих кібератак. Проте впровадження алгоритмів МН у контури безпеки реального часу стикається зі значними перешкодами, що зумовлені математичною природою моделей та фізичними обмеженнями обчислювальних ресурсів мережі.

Основний матеріал. Головною проблемою застосування моделей машинного навчання для безперервного захисту КІС є суперечність між

точністю виявлення загроз та швидкістю прийняття рішень. Сучасна мережева інфраструктура генерує колосальні масиви багатовимірних даних (логи, мережеві пакети, телеметрія поведінки). Навчання та інференс на таких обсягах даних призводять до так званого «прокляття розмірності» [2].

З математичної точки зору, завдання оптимізації адаптивної системи захисту зводиться до мінімізації зваженої функції втрат $L(\theta)$, яка враховує помилки першого (хибне спрацьовування, False Positives) та другого (пропуск загрози, False Negatives) роду:

$$\min_{\theta}\left(\sum_{i=1}^{N} L\left(y_i, f(x_i;\theta)\right) + \lambda R(\theta)\right), \quad (1)$$

де θ — параметри моделі, $R(\theta)$ — регуляризаційний член для запобігання перенавчанню, за умови жорсткого обмеження на час інференсу $t_{inf}(\theta) \leq T_{max}$.

Для вирішення цієї задачі необхідно застосовувати методи зменшення розмірності ознакового простору (наприклад, метод головних компонент (PCA) або автоенкодери), що дозволяє знизити обчислювальну складність алгоритму без критичної втрати інформативності [3]. Іншою проблемою є сильний дисбаланс класів: легітимний трафік становить понад 99%, тоді як аномалії — менше 1%. Це вимагає застосування спеціалізованих метрик оптимізації (F1-score, Precision-Recall AUC) замість стандартної точності (Accuracy), а також синтетичного балансування даних (наприклад, алгоритм SMOTE) або використання алгоритмів ізоляції аномалій (Isolation Forest).

Висновки

Впровадження інтелектуальних систем безперервного захисту в динамічних КІС неможливе без глибокої математичної оптимізації алгоритмів машинного навчання. Забезпечення стійкості до кібератак у реальному часі вимагає створення полегшених, ресурсоефективних ансамблевих моделей, здатних мінімізувати хибні спрацьовування в умовах високорозмірного та незбалансованого потоку даних. Подальші дослідження доцільно спрямувати на розробку градієнтних методів оптимізації гіперпараметрів моделей спеціально для архітектури Zero Trust.

ЛІТЕРАТУРА

1. Rose S., Borchert O., Mitchell S., Connelly S. Zero Trust Architecture. NIST Special Publication 800-207. 2020. 53 p. URL: https://doi.org/10.6028/NIST.SP.800-207
2. Ахметов Б. С., Корченко О. Г. Моделювання систем захисту інформації: сучасні підходи та алгоритми. Захист інформації. 2022. Т. 24, № 1. С. 15–24.
3. Apruzzese G., Colajanni M., Ferretti L. Evaluating the effectiveness of Machine Learning for cyber security. IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). 2018. P. 1–6.

PROBLEMS OF OPTIMIZING MACHINE LEARNING MODELS FOR CONTINUOUS PROTECTION OF CORPORATE INFORMATION SYSTEMS

Volodymyr Riabenko

Abstract. *The work investigates the mathematical and algorithmic problems of optimizing machine learning (ML) models used for continuous monitoring and protection of modern corporate information systems (CIS). The growth of network traffic volumes and the transition to the Zero Trust architecture require security systems to analyze data in real-time, leading to a critical increase in computational load. The key barriers to ML implementation are considered, in particular, the high dimensionality of the feature space, the problem of class imbalance, and latency during inference. The necessity of applying dimensionality reduction methods and ensemble approaches to increase the accuracy of detecting multi-vector cyberattacks while reducing the false positive rate is justified. Directions of mathematical modeling for loss function minimization under dynamic changes in CIS are proposed.*

Keywords: *corporate information systems, machine learning, algorithm optimization, continuous protection, Zero Trust architecture, computational complexity.*

REFERENCE

1. Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). Zero Trust Architecture. NIST Special Publication 800-207, 53. https://doi.org/10.6028/NIST.SP.800-207
2. Akhmetov, B. S., & Korchenko, O. H. (2022). Modeliuvannia system zakhystu informatsii: suchasni pidkhody ta algorytmy [Modeling of information security systems: modern approaches and algorithms]. Zakhyst informatsii, 24(1), 15-24 [in Ukrainian].
3. Apruzzese, G., Colajanni, M., & Ferretti, L. (2018). Evaluating the effectiveness of Machine Learning for cyber security. IEEE/ACM 26th International Symposium on Quality of Service (IWQoS), 1-6.