УДК 004.056.5:004.852:004.891.2:004.272

Фесенко Тетяна Миколаївна
кандидат технічних наук, доцент,
доцент кафедри комп'ютерних та інформаційних технологій і систем
Національний університет «Полтавська політехніка ім. Ю. Кондратюка»
Калашнікова Юлія Вадимівна
асистент кафедри комп'ютерних та інформаційних технологій і систем
Національний університет «Полтавська політехніка ім. Ю. Кондратюка»

МОДЕЛЬ ПРОГНОЗУ ЗАГРОЗ У ZERO TRUST-АРХІТЕКТУРІ

У роботі представлено комплексний методологічний підхід до проєктування інтелектуальної системи прогнозування компрометації облікових записів у корпоративних інформаційних середовищах. Запропонована модель (Рис.1) інтегрує федеративне навчання, графові нейронні мережі та пояснювальний штучний інтелект у рамках парадигми Zero Trust Architecture [1]. Така інтеграція забезпечує підвищену стійкість системи автентифікації та управління доступом завдяки децентралізованій обробці даних, збереженню конфіденційності користувачів і можливості адаптивного навчання моделей без централізованої агрегації журналів безпеки.

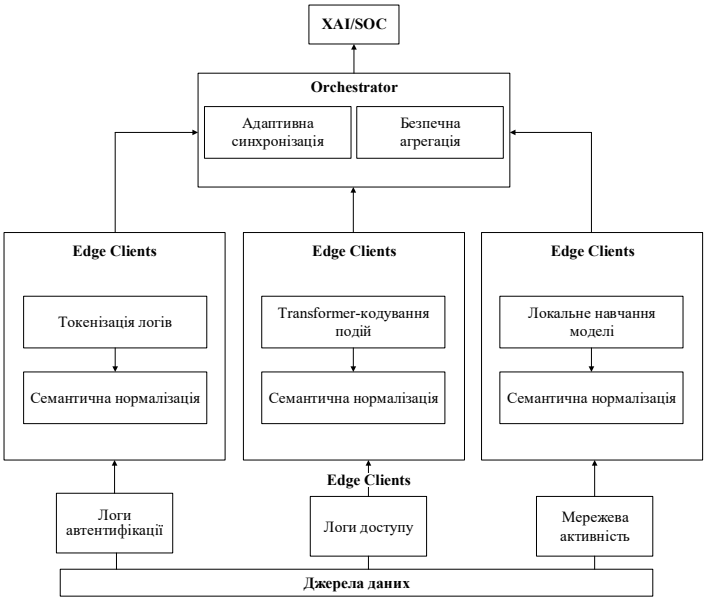

Рис. 1. Модель федеративної системи прогнозування компрометації облікових записів користувачів

Однією з ключових відмінностей запропонованого підходу є реалізація локального навчання на вузлах корпоративної інфраструктури без передачі первинних даних до центрального сховища. Агрегація параметрів здійснюється за допомогою безпечних протоколів федеративної системи, що враховують гетерогенність і асинхронність клієнтів. Графовий модуль

формує міжкористувацькі та міжсистемні зв'язки у вигляді динамічного орієнтованого графа [2], у якому вузли представляють сутності користувачів і ресурсів, а ребра – події доступу або поведінкові залежності. Це дозволяє ідентифікувати приховані аномальні патерни та ризики компрометації на рівні міжсистемних взаємодій.

Для підвищення прозорості прогнозів у систему інтегровано компонент пояснювального ШІ, який генерує причинно-наслідкові інтерпретації та контрфактичні пояснення рішень моделі. XAI-модуль функціонує у вигляді автономного мікросервісу з API-взаємодією через шину повідомлень, забезпечуючи ізольовану обробку даних і можливість побудови «why-графів», які пояснюють, чому конкретна поведінка користувача була класифікована як ризикова. На глобальному рівні пояснення агрегуються без передачі сирих даних за допомогою federated XAI aggregator [3], що зберігає конфіденційність клієнтів і формує інтегровані вектори пояснень для SOC-аналітики.

З технічного боку архітектура реалізує асинхронну стратегію оновлення параметрів, що дозволяє системі реагувати на зміни поведінкових патернів користувачів у режимі реального часу. Для компенсації запізнілих оновлень застосовується функція синхронізації із часовими вагами, а для збереження попередніх станів моделі – регуляризаційний механізм Elastic Weight Consolidation, який мінімізує катастрофічне забування. Конфіденційність гарантовано механізмами secure aggregation та диференційного шуму, що унеможливлюють відновлення первинних даних користувачів під час навчання.

Таким чином, розроблений підхід формує науково обґрунтовану платформу для створення інтелектуальних систем кіберзахисту нового покоління, у яких алгоритмічна прозорість поєднується з конфіденційністю та масштабованістю. Це відкриває перспективи подальшого розвитку у напрямах Federated Reinforcement Learning, енергоефективного обчислення на периферії та інтеграції квантово-стійких протоколів безпечної агрегації.

Список використаних джерел

1. Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero Trust architecture* (NIST Special Publication 800-207). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-207.

2. Kim, H., Lee, B. S., Shin, W.-Y., & Lim, S. (2022). *Graph Anomaly Detection With Graph Neural Networks: Current Status and Challenges*. IEEE Access, 10, 111820–111829. https://doi.org/10.1109/ACCESS.2022.3211306.

3. Gupta, L., & Misra, D. C. (2025). *Cybersecurity Threat Detection Through Explainable Artificial Intelligence (XAI): A Data-Driven Framework.* International Research Journal of MMC, 6 (2), 119–131. https://doi.org/10.3126/irjmmc.v6i2.80687.