ВИКОРИСТАННЯ АІ ТА МАШИННОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ АНОМАЛІЙ У МЕРЕЖАХ ІЗ НУЛЬОВОЮ ДОВІРОЮ

Куля В. М.
Харківський національний університет радіоелектроніки, Харків, Україна

Сучасна парадигма забезпечення кібербезпеки зазнає фундаментальних трансформацій через неефективність традиційних моделей захисту периметра в умовах інтенсивного впровадження хмарних обчислень та віддаленого доступу. Архітектура нульової довіри (Zero Trust Architecture, ZTA), яка базується на повній відсутності апріорної довіри до будь-якого суб'єкта, потребує миттєвого аналізу величезних масивів даних. Статичні правила доступу не здатні адаптуватися до динамічних змін у ландшафті загроз, що зумовлює критичну необхідність інтеграції штучного інтелекту та машинного навчання для забезпечення безперервної інтелектуальної верифікації.

Метою доповіді є обґрунтування ролі та визначення ключових напрямів застосування методів машинного навчання для виявлення аномалій у мережевих структурах, побудованих на принципах нульової довіри, а також аналіз механізмів трансформації статичних моделей безпеки в адаптивні системи предиктивного захисту.

Реалізація інтелектуальних систем у межах архітектури нульової довіри (Zero Trust Architecture, ZTA) вимагає концептуального переходу від традиційних статичних правил контролю доступу до динамічного механізму безперервного оцінювання довіри (Trust Engine) [1, 2]. В основі цього процесу лежить глибока інтеграція інструментів аналізу поведінки користувачів та сутностей (User and Entity Behavior Analytics, UEBA), які акумулюють розширену мережеву телеметрію для формування багатовимірних профілів нормальної активності. Застосування алгоритмів глибокого навчання без учителя (Unsupervised Deep Learning), зокрема автоенкодерів (Autoencoders), забезпечує високоточну ідентифікацію латентних відхилень шляхом обчислення помилок реконструкції даних, що дозволяє виявляти атаки нульового дня та компрометацію облікових даних [3]. Для аналізу часових і послідовних залежностей, які є характерними для складних цілеспрямованих загроз (APT) та етапів горизонтального переміщення (Lateral Movement), архітектура ефективно доповнюється рекурентними моделями, такими як двонаправлені мережі довгої короткострокової пам'яті (Bi-LSTM) [1, 3]. Комплексний аналіз цих поведінкових патернів дає змогу ML-моделям автоматично обчислювати динамічний коефіцієнт ризику (Risk Scoring) у реальному часі [1]. Зпираючись на цей показник, точка прийняття рішень щодо політик (Policy Decision Point, PDP) генерує контекстно-залежні інструкції для точки застосування політик (Policy Enforcement Point, PEP) [1]. Залежно від виявленого рівня загрози, система здатна автономно ініціювати посилену багатофакторну автентифікацію (MFA), обмежити привілеї користувача або виконати превентивну мікросегментацію для повної ізоляції скомпрометованого вузла.

Разом із тим, імплементація методів штучного інтелекту в критичні інфраструктури неминуче стикається з проблемою вразливості самих ML-моделей до змагальних атак (Adversarial Attacks), зокрема цілеспрямованого «отруєння даних» (Data Poisoning) під час навчання та ухилення від виявлення. Для підвищення загальної кіберстійкості архітектури доцільним є впровадження байєсівських нейронних мереж, які запобігають прийняттю системою надмірно впевнених хибних рішень під час реєстрації рідкісних або неоднозначних подій. Окрім того, фундаментальною вимогою залишається подолання проблеми алгоритмічного «чорного ящика» за допомогою інтеграції методів пояснюваного штучного інтелекту (Explainable AI, XAI), зокрема алгоритмів SHAP або LIME [3]. Це перетворює приховані процеси прийняття рішень на зрозумілі візуальні моделі, забезпечуючи інтерпретованість результатів, підтримуючи концепцію управління за участю людини та гарантуючи довіру до автономних предиктивних дій системи. Синергія принципів архітектури нульової довіри (Zero Trust Architecture, ZTA) та передових технологій машинного навчання (ML) є фундаментальним фактором формування резильєнтного кіберпростору, що остаточно нівелює обмеження традиційних статичних периметрових моделей. Інтеграція інтелектуальних інструментів, зокрема аналізу поведінки користувачів та сутностей (UEBA) на базі глибокого навчання, перетворює систему захисту на проактивну «імунну систему» мережі, здатну здійснювати безперервну автентифікацію та динамічно обчислювати оцінку ризику (Risk Scoring) у режимі реального часу [1]. Цей підхід забезпечує автоматизоване застосування політик безпеки, що мінімізує вплив людського фактора, критично зменшує площу атаки та ефективно блокує горизонтальне переміщення (Lateral Movement) зловмисників у багатохмарних середовищах шляхом превентивної мікросегментації [4]. Водночас надійність таких адаптивних систем безпосередньо залежить від їхньої захищеності від змагальних атак та цілеспрямованого отруєння даних, що вимагає застосування стійких імовірнісних моделей, таких як байєсівські нейронні мережі (BNN), для врахування прогностичної невизначеності середовища.

Список літератури

1. Sivanageswara Rao Gandikota. Deep learning-enabled zero trust architecture for intelligent identity and access management in cloud ecosystems 2024 DOI: https://doi.org/10.30574/wjarr.2024.22.3.1842

2. Moskvin K., Sievierinov O. Zero Trust Architecture in Corporate Cybersecurity Systems // Computer and information systems and technologies : Seventh International Scientific and Technical Conference, 2024. – Kharkiv : NURE, 2024. – p. 54-55.

3. Maria Gabriela Camila. A Multi-Tenant Cloud Security Framework Using Zero-Trust Architecture and AI-Based Anomaly Detection 2024 DOI:10.63282/3117-5481/AIJCST-V6I4P101

4. Ravi Gupta and Guneet Bhatia. Zero-trust industrial network security using AI and explainable inference novelty. World Journal of Advanced Research and Reviews, 2025. DOI: https://doi.org/10.30574/wjarr.2025.28.2.3705