DOI: https://doi.org/10.15276/ict.02.2025.23
УДК 004.056.55:004.451.85

Керована ін'єкція аномалій у мережевий трафік для stress-тестування систем виявлення вторгнень

Науменко Роман Іванович[1)]
Аспірант каф. Комп'ютерних інтелектуальних систем та мереж
ORCID: https://orcid.org/0009-0003-7612-1773; lwoodooz7@gmail.com
Буюклі Віктор Сергійович[1)]
Аспірант каф. Комп'ютерних інтелектуальних систем та мереж
ORCID: https://orcid.org/0000-0001-7384-2290; vktr.buyukli@gmail.com
Тішин Петро Метталинович[1)]
Канд. фіз.-матем. наук, доцент каф. Комп'ютерних інтелектуальних систем та мереж
ORCID: https://orcid.org/0000-0003-2506-5348; petrmettal@gmail.com. Scopus Author ID: 57190400970
Мартинюк Олександр Миколайович[1)]
Канд. техніч. наук, доцент, завідувач каф. Комп'ютерних інтелектуальних систем та мереж
ORCID:http://orcid.org/0000-0003-1461-2000; martynyuk@op.edu.ua. Scopus Author ID: 57103391900
[1)] Національний університет «Одеська політехніка», пр. Шевченка, 1. Одеса, 65044, Україна

АНОТАЦІЯ

В умовах сучасних динамічних мереж, де шифрування трафіку та використання витончених «low-and-slow» атак стають нормою, класичні підходи до тестування систем виявлення вторгнень (IDS) демонструють свою неспроможність. Існуючі статичні набори даних не відображають ані сучасних багатоетапних сценаріїв атак, ані різноманіття легітимного фонового трафіку, що призводить до нерепрезентативних оцінок. У цій роботі пропонується комплексна методологія керованої ін'єкції аномалій, призначена для систематичного та відтворюваного stress-тестування IDS в лабораторних умовах. На відміну від застарілих практик, запропонований фреймворк вводить чітко визначені та керовані параметри ін'єкцій. Це дозволяє гнучко задавати інтенсивність, тривалість, просторово-часову локальність, рівень непомітності та семантику сценаріїв, що охоплюють розвідку, атаки на відмову в обслуговуванні, перебір облікових даних та приховану ексфільтрацію даних. Такий підхід уможливлює цілеспрямоване «налаштування» складності перевірок і прозоре порівняння різних реалізацій IDS. Ключовою перевагою є повна незалежність від конкретної технології синтезу: методологія сумісна з різними генераторами, зокрема із сучасними дифузійними моделями. Це гарантує довгострокову актуальність фреймворку, дозволяючи йому еволюціонувати разом із розвитком генеративних технологій. Ін'єкції інтегруються в наявні трейси або формуються разом із реалістичним фоном, зберігаючи гарантовану відтворюваність через стандартизовані описові маніфести. Ці маніфести фіксують версії інструментів, початкові значення генераторів випадковості та контрольні суми артефактів. Таким чином, стандартизуються не лише умови атаки, але й протокол оцінювання. Планується проведення пілотних випробувань, які, як очікується, продемонструють контрольований вплив параметрів на поведінку IDS. Очікується, що зі зростанням непомітності сигнатурні детектори втрачатимуть ефективність, тоді як поведінкові демонструватимуть вимірюване збільшення часу реакції. Це дозволить досліджувати їхні граничні режими роботи, аналізувати рівень хибних спрацювань та оцінювати ресурсну поведінку систем під навантаженням. Отримані спостереження мають підтвердити придатність методології для планомірної оцінки стійкості та виявлення «сліпих зон». Окремо розглянуто безпекові та етичні аспекти. У якості подальших робіт окреслено інтеграцію фреймворку з процесами CI/CD, що сприятиме впровадженню культури безперервної валідації захисту (DevSecOps), та публікацію відкритих, відтворюваних бенчмарків

Ключові слова: керована ін'єкція аномалій; мережевий трафік; stress-тестування; IDS; низькопомітні атаки; DevSecOps

Актуальність. Сучасні IDS працюють у динамічних, зашифрованих та різнорідних середовищах, де класичні датасети та випадкові «аномальні» вставки не відтворюють реальні сценарії загроз [3]. Навіть відносно нові набори даних, такі як CIC-IDS-2017, мають свої обмеження у відтворенні сучасних атак. Через це лабораторні результати погано корелюють із поведінкою систем у реальних умовах: системи переоцінюють чутливість або пропускають «повільні» й малопомітні атаки, а також залишаються вразливими до adversarial-атак [5].

Керована ін'єкція аномалій дає відтворюваний спосіб формувати події з контрольованою складністю, непомітністю та семантикою [7]. Це робить можливим систематичне stress-тестування й прозоре порівняння різних IDS за однакових умов. Підхід природно інтегрується у CI/CD, забезпечуючи регулярні регресійні перевірки безпеки. На тлі

This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/deed.uk)

тотального шифрування, коли зміст пакетів недоступний і зростає роль поведінкових ознак, що аналізуються такими системами як Zeek/Bro [1], контрольоване моделювання стає критично важливим. Це зменшує залежність від застарілих датасетів, сприяє відкритим бенчмаркам і підвищує довіру до результатів, що робить методологію практично необхідною для сучасного кіберзахисту.

Метою роботи є розробка та оцінка підходу керованої ін’єкції аномалій у мережевий трафік для відтворюваного та систематичного stress-тестування систем виявлення вторгнень (IDS). Запропонована методологія передбачає введення керованих параметрів аномалій (інтенсивність, тривалість, локальність, рівень непомітності та семантика), що покликано підвищити достовірність і репрезентативність випробувань у порівнянні з традиційними наборами даних [3].

Важливою частиною дослідження є перевірка здатності підходу узагальнюватися на різні профілі трафіку, протоколи та інструменти генерації, не використані під час налаштування сценаріїв. Таким чином, метою є створення методології, яка буде не лише точнішою у відображенні реальних умов і «повільних»/малопомітних атак, але й більш універсальною – придатною до інтеграції у різні середовища та інженерні процеси (зокрема CI/CD) для регулярної оцінки стійкості IDS.

Основна частина. Оцінка сучасних систем виявлення вторгнень (IDS) часто спирається на фіксовані датасети, що погано відтворюють реальні умови, такі як шифрування трафіку та «повільні» атаки [3, 6]. Як наслідок, лабораторні результати слабо корелюють із поведінкою систем у продакшені, а висновки щодо їхньої стійкості виявляються нестабільними. У відповідь на цю проблему в роботі пропонується підхід керованої ін'єкції аномалій, що забезпечує відтворювані випробування та систематичний стрес-тест IDS за контрольованих умов [7]. Його ключова ідея полягає у заданні чітких параметрів ін'єкцій (інтенсивність, локальність, непомітність, семантика, тривалість) для отримання прогнозованої зміни складності для детектора. Це дозволяє прозоро порівнювати різні реалізації IDS, виявляти їх «сліпі зони» та перевіряти граничні режими роботи.

В основі методології лежить розділення «що» ін'єктується (семантика події, наприклад, TCP-скан чи DNS-тунелювання) та «як» це відбувається (керовані параметри). Такий поділ забезпечує незалежність від технології синтезу, дозволяючи використовувати будь-які генератори – від емуляторів до сучасних генеративних моделей [4]. Для забезпечення відтворюваності кожна ін'єкція супроводжується маніфестом, що фіксує всі параметри, версії інструментів та джерела випадковості, усуваючи «тонкий підгін» [7]. Особлива увага приділяється параметру непомітності, який дозволяє моделювати складні для виявлення “low-and-slow” атаки шляхом узгодження статистичних характеристик аномалії з фоновим трафіком. Це є справжнім викликом для сучасних поведінкових [1] та гібридних [5] детекторів.

Практична реалізація є послідовним конвеєром: від профілювання фонового трафіку до конфігурації ін'єкції через маніфест, її накладання та семантичної валідації. Результатом є готовий, маркований датасет. План оцінювання використовує ці датасети для контрольованого тестування за правилом «одна зміна — один запуск» з кількома повторами. Для кожного сценарію створюється градація складності, що дозволяє порівнювати реакцію різних IDS за ключовими метриками: успішність виявлення, час реакції та рівень хибних спрацювань.

Оглядовий каталог сценаріїв ін'єкцій, що включає TCP SYN-сканування, HTTP brute-force, імпульсний DoS, DNS-тунелювання, low-and-slow ексфільтрацію та періодичний C2-зв'язок, детально описаний у Таблиці. У ній наведено контрольовані параметри та обмеження семантики для кожного типу атаки.

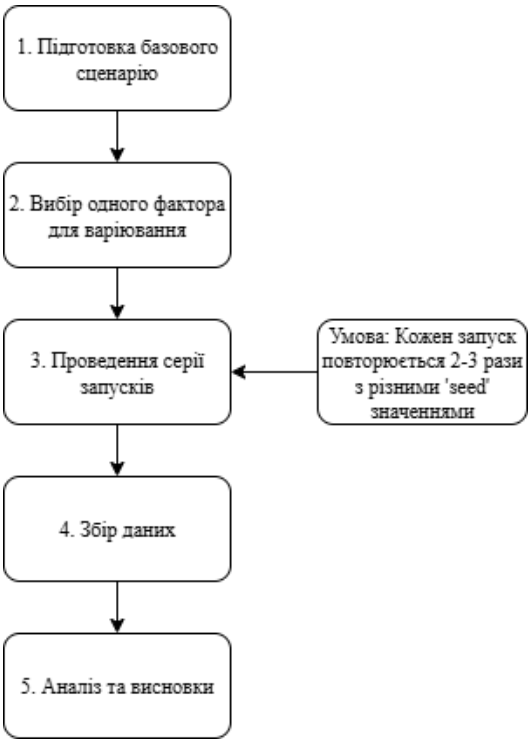

Рис. 1. **Схема дизайну експерименту, що базується на правилі «одна зміна – один запуск» з кількома повторами для забезпечення статистичної значущості**

Таблиця. **Сценарії ін'єкцій та контрольовані параметри**

Сценарій	Рівень	Інтенсивність	Тривалість	Локальність	Непомітність	Коротка примітка
TCP SYN-сканування	Потік	Від розріджених проб до суцільного перебору	-	Одна підмережа або весь сегмент	Середня (стохастичні паузи, «міксування» траєкторій)	Класичний сценарій розвідки.
HTTP brute-force	Сесія	5-60 запитів/хв	Кілька хвилин - година	Один вузол	Висока (імітація людського ритму, джитери, різні user-agent)	Перевірка стійкості до атак на облікові дані.
Імпульсний DoS	Потік	Burst-навантаження зі змінною амплітудою	30-120 секунд	Сегмент мережі	Низька (але з регульованою періодичністю імпульсів)	Тестування реакції на короткочасне перевантаження.
DNS-тунелювання	Сесія	0.05-0.5 запитів/с	Десятки хвилин	Один вузол	Висока (варіації ентропії в іменах, TTL, часових інтервалів)	Сценарій прихованої передачі даних.
"Low-and-slow" ексфільтрація	Сесія/ Потік	Мінімальна (дрібні порції з тривалими паузами)	-	Один вузол	Дуже висока (підгонка розмірів і таймінгу під фон сервісу)	Імітація витончених APT-загроз.
Періодичний C2-зв'язок	Сесія	Рідкісні «пінги» з невеликим джитером у таймінгу	-	Один вузол	Середня-Висока	Моделювання активності бекдора/трояна.

Пілотне дослідження. Для валідації підходу планується провести пілотне дослідження на реалістичному фоні з двома репрезентативними сценаріями: розрідженим TCP SYN-скануванням та low-and-slow ексфільтрацією. Планується порівняти сигнатурний двигун

(наприклад, на базі Snort [2]) та поведінковий детектор (наприклад, на базі Zeek/Bro [1]) на трьох рівнях складності для кожного сценарію.

Очікується побачити передбачувану залежність: зі зростанням непомітності має зростати час до виявлення та збільшуватися невизначеність спрацьовувань. Водночас підвищення інтенсивності, ймовірно, покращить виявлення у сигнатурних системах, але не завжди допоможе поведінковим. У випадку ексфільтрації очікується, що поведінковий детектор буде чутливішим до варіацій таймінгу, але втрачатиме стабільність при підгонці інтервалів під фонові ритми, тоді як сигнатурна система слабко реагуватиме на такі події. Таким чином, очікується, що пілотне дослідження підтвердить корисність керованих ін'єкцій для відтворення граничних умов та прозорого порівняння різних підходів.

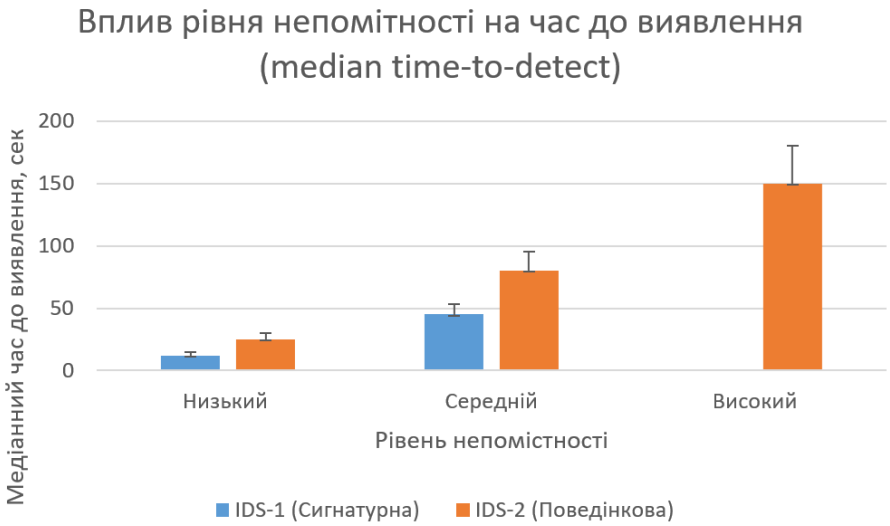

Рис. 2 . **Вплив рівня непомітності атаки на медіанний час до виявлення (median time-to-detect) для сигнатурної (IDS-1) та поведінкової (IDS-2) систем**

На графіку наведено умовні дані, що ілюструють очікувану залежність. «Вуса» на діаграмі позначають 95 % довірчий інтервал.

Обмеження та практичні рекомендації. Ключові обмеження підходу полягають у неминучому розриві між синтетичними та реальними атаками, а також у ризику переадаптації IDS до тестових сценаріїв [6]. Для мінімізації цих ризиків рекомендується практикувати ротацію генераторів, перехресні перевірки на різних джерелах фону та проводити семантичну валідацію трафіку. Практичні рекомендації для впровадження зводяться до інтеграції тестування у CI/CD процеси: щоденне виконання «легких» сценаріїв, «середніх» – щотижня перед релізами, а «складних» – за зміни політик IDS.

Відтворюваність, безпека та етика. Запропонований підхід забезпечує високий рівень відтворюваності завдяки супроводженню кожної ін'єкції маніфестом, що фіксує всі параметри експерименту [8]. З етичних міркувань та міркувань безпеки, всі випробування проводяться виключно на ізольованих стендах з деперсоналізованими даними, а доступ до шкідливих артефактів суворо обмежується лабораторним середовищем.

Висновки. Запропонована рамка керованої ін'єкції аномалій дозволяє відтворювано й системно перевіряти стійкість IDS у реалістичних умовах, регулюючи складність подій без прив’язки до конкретного генератора. Стандартизований протокол та каталог сценаріїв роблять результати міжкомандно співставними, а плановане пілотне дослідження має підтвердити практичність підходу для виявлення «сліпих зон». Обмеження синтетики компенсуються семантичною валідацією та дисципліною маніфестів. З огляду на тренди шифрування та еволюцію атак, інтеграція такого тестування в інженерні процеси є доцільною як на рівні R&D, так і для регулярного контролю якості захисту.

СПИСОК ЛІТЕРАТУРИ

1. Paxson V. “Bro: A system for detecting network intruders in real-time”. *Computer Networks,* 1999; 31 (23-24): 2435–2463. DOI: https://doi.org/10.1016/S1389-1286(99)00112-7.

2. Roesch M. “Snort: lightweight intrusion detection for networks”. *Proceedings of the 13th USENIX conference on System Administration (LISA '99)*. 1999. p. 229–238.

3. Ring M., et al. “A survey of network-based intrusion detection data sets”. *Computers & Security*. 2019; 86: 147–167. DOI: https://doi.org/10.1016/j.cose.2019.06.005.

4. Gong X., Chen S., Li N. “A network traffic data generation model based on AOT-DDPM for abnormal traffic detection”. *Evolving Systems.* 2025; 16: 15. DOI: https://doi.org/10.1007/s12530-024-09644-y.

5. Buczak A. L., Guven E. “A survey of data mining and machine learning methods for cyber security intrusion detection”. *IEEE Communications Surveys & Tutorials*. 2016; 18 (2): 1153–1176. DOI: https://doi.org/10.1109/COMST.2015.2494502.

6. Buchta R., Gkoktsis G., Kleiner C. “Advanced persistent threat attack detection systems: a review of approaches, challenges, and trends”. *Digital Threats: Research and Practice*. 2024; 5 (4): 1–37. DOI: https://doi.org/10.1145/3696014.

DOI: https://doi.org/10.15276/ict.02.2025.23
UDC 004.056.55:004.451.85

Controlled anomaly injection into network traffic for stress-testing of intrusion detection systems

Roman I. Naumenko[1)]
PhD Student of the Department of Intellectual Systems and Networks
ORCID: https://orcid.org/0009-0003-7612-1773; lwoodooz7@gmail.com
Viktor S. Buiukli[1)]
PhD Student of the Department of Computer Intellectual Systems and Networks
ORCID: https://orcid.org/0000-0001-7384-2290; vktr.buyukli@gmail.com
Petr M. Tishyn[1)]
PhD, Associate Professor of the Department of Computer Intellectual Systems and Networks
ORCID: https://orcid.org/0000-0003-2506-5348; petrmettal@gmail.com. Scopus Author ID: 57190400970
Oleksandr N. Martynyuk[1)]
PhD, Associate Professor of the Department of Computer Intellectual Systems and Networks
ORCID: http://orcid.org/0000-0003-1461-2000; martynyuk@op.edu.ua. Scopus Author ID: 57103391900
[1)] Odesa Polytechnic National University, 1, Shevchenko Ave. Odesa, 65044, Ukraine

ABSTRACT

In modern dynamic networks where traffic encryption and sophisticated “low-and-slow” attacks are becoming the norm, classical approaches to testing Intrusion Detection Systems (IDS) are proving to be inadequate. Existing static datasets fail to reflect either modern multi-stage attack scenarios or the diversity of legitimate background traffic, leading to unrepresentative evaluations. This paper proposes a comprehensive methodology of controlled anomaly injection for the systematic and reproducible stress-testing of IDS in a laboratory environment. In contrast to outdated practices, the proposed framework introduces clearly defined and controlled injection parameters. This allows for flexible configuration of the intensity, duration, spatio-temporal locality, stealth level, and semantics of scenarios, covering reconnaissance, denial-of-service attacks, credential brute-forcing, and covert data exfiltration.Such an approach enables the targeted "tuning" of test complexity and the transparent comparison of different IDS implementations. A key advantage is its complete independence from any specific synthesis technology; the methodology is compatible with various generators, including modern diffusion models. This ensures the long-term relevance of the framework, allowing it to evolve alongside the development of generative technologies. Injections are integrated into existing traces or co-generated with a realistic background, maintaining guaranteed reproducibility through standardized descriptive manifests. These manifests record tool versions, random generator seeds, and artifact checksums. Thus, not only the attack conditions but also the evaluation protocol are standardized.Pilot tests are planned and expected to demonstrate the controlled impact of the parameters on IDS behavior. It is anticipated that with an increase in stealth, signature-based detectors will lose effectiveness, while behavioral detectors will show a measurable increase in reaction time. This will make it possible to investigate their operational limits, analyze false positive rates, and evaluate the resource behavior of the systems under load. The findings should confirm the methodology's suitability for the systematic assessment of resilience and the identification of “blind spots.” Security and ethical aspects are considered separately. Future work outlines the integration of the framework with CI/CD processes to foster a culture of continuous security validation (DevSecOps), and the publication of open, reproducible benchmarks.

Keywords: controlled anomaly injection; network traffic; stress testing; IDS; low-and-slow attacks; DevSecOps