

**Редзюк Євгеній Володимирович**

кандидат технічних наук,

начальник кафедри комп'ютерних наук та інтелектуальних технологій

Військовий інститут телекомунікацій та інформатизації імені Героїв Крут, Київ, Україна

ORCID: 0000-0001-5592-5121

yevhenii.redziuk@viti.edu.ua

**ФАКТОРНА ДІАГНОСТИКА ДИСКРИМІНАТИВНОЇ ЗДАТНОСТІ ЛОКАЛЬНИХ ГРАФОВИХ
ОЗНАК У ЗАДАЧІ ВИЯВЛЕННЯ АНОМАЛІЙ НА ТЕМПОРАЛЬНИХ МУЛЬТИГРАФАХ**

Анотація. Виявлення аномалій на темпоральних графах взаємодії є задачею інтелектуального аналізу даних, у якій потокоорієнтовані методи обмежено враховують структурний контекст взаємодій між вузлами. Метою роботи є оцінювання дискримінативної здатності локальних інтерпретованих графових ознак як самостійного детектора мережових аномалій за некерованої побудови ознак. Запропоновано композитну метрику аномальності на основі трьох графових компонентів: відхилення центральності за посередництвом, зміни спільноти Louvain та Жаккардового розходження локальної околиці. Експериментальну оцінку проведено на наборі CICIDS2017 із використанням контрольного дизайну «класифікатор × ознаковий простір» та порівнянням із потокоорієнтованими базовими методами. Метод реалізовано на стеку Neo4j Graph Data Science та Java для аналізу темпорального мультиграфа мережевого трафіку. Результати показали, що досліджені графові ознаки зберігають ранжувальну здатність (ROC-AUC до 0,84), але не забезпечують операційно прийнятної точності як самостійний детектор: у спільному протоколі оцінювання $F_1 = 0,0021$ проти $F_1 = 0,2000$ у потокового базового методу. Контрольні експерименти локалізували джерело деградації в ознаковому просторі, а не у класифікаторі, кількості ознак чи некерованому характері побудови ознак. Наукова новизна роботи полягає у встановленні обмежень дискримінативної здатності локальних графових ознак у досліджуваній постановці задачі та у розробленні діагностичного шаблону для аудиту графових детекторів аномалій, що поєднує декомпозицію розмітки, контрольоване порівняння класифікаторів і ознакових просторів та аналіз структурних і семантичних чинників низької результативності. Результати обґрунтовують доцільність використання потокових ознак як основи виявлення аномалій у досліджуваному класі корпоративних мереж; узагальнення висновків на сітчасті топології потребує окремої експериментальної перевірки.

Ключові слова: інтелектуальний аналіз даних; графова аналітика; виявлення аномалій; темпоральний мультиграф; графові ознаки; виявлення спільнот; некероване навчання; виявлення вторгнень.

ВСТУП

Сучасні мережі оперують обсягами трафіку, які роблять ручний моніторинг неможливим. Системи виявлення вторгнень (Intrusion Detection Systems, IDS) мають виявляти поведінкові відхилення, характерні для цілеспрямованих атак, прихованих сканувань, бокового переміщення та екфільтрації даних [1] – [3]. Більшість наявних IDS-рішень побудована на парадигмі аналізу потоків (потокоорієнтована парадигма), де одиницею аналізу є окремий мережовий потік, описаний вектором статистичних ознак. Така парадигма має три принципові обмеження:

- ігнорування контексту взаємодій – структурна аномалія є властивістю сукупності потоків і зв'язків між вузлами, а не окремого потоку;
- слабе виявлення повільних атак (повільне сканування, повільний перебір паролів, дозовані екфільтрації), свідомо спроектованих залишатися під порогами потокоорієнтованих детекторів;
- нездатність виявляти раптову зміну ролі вузла – перетворення робочої станції на вузол-плацдарм з високою центральністю у графі комунікацій [1], [3].

Перехід до графової парадигми усуває ці обмеження: трафік представляється як темпоральний граф, а виявлення аномалій – як застосування мір центральності, виявлення спільнот, аналізу подібності околиць [1], [3], [4]. Промислове впровадження стримується складністю інфраструктури, необхідної для



використання Python-бібліотек машинного навчання, а також обчислювальною складністю графових нейронних мереж (Graph Neural Networks, GNN), що може ускладнювати їх розгортання в умовах обмежених ресурсів та забезпечення інтерпретованості результатів [2].

Постановка проблеми. Нехай мережева взаємодія подана темпоральним мультиграфом, вузли якого відповідають хостам, а ребра – потокам із часовими мітками. Для аналітичного вікна фіксованої тривалості кожному хосту зіставляється вектор локальних графових ознак, обчислених без міток атак: відхилення центральності за посередництвом від базового профілю, зміна приналежності до спільноти та розходження локальної околиці. Еталонна розмітка задана на рівні пари «хост × вікно», частка позитивних пар не перевищує 0,15 %.

Потрібно визначити, чи утворюють такі ознаки ознаковий простір, достатній для виявлення аномальних хостів на рівні потокоорієнтованих базових методів, а якщо ні – локалізувати причину серед чотирьох чинників: парадигма представлення даних, класифікатор, природа ознак і їх кількість. Обмеження: на тренувальному сегменті дозволено калібрувати лише поріг і ваги композиції; обчислення виконуються класичними графовими алгоритмами на промисловому стеку без GNN. Нетривіальність задачі в тому, що позитивні результати для локальних графових ознак отримано на сітчастих (mesh) топологіях і переважно засобами GNN, тоді як для зіркоподібних корпоративних топологій систематичної діагностики немає, а негативний результат без атрибуції причини не має методологічної цінності.

Аналіз останніх досліджень і публікацій. Виявлення мережевих вторгнень історично формулювалося як класифікація окремих потоків [2]. Структурна аномалія на рівні зв'язків між вузлами (боковий рух, координовані атаки, екфільтрація) для таких методів залишається невидимою. Напрям графового виявлення аномалій представлений таксономією [4] та оглядом глибокого навчання на графах [3].

Графові методи поділяються на класичні алгоритми (центральність, виявлення спільнот, сигнатури підграфів) і GNN-підходи [2], [3]. Класичні забезпечують інтерпретовність – оператор бачить причину аномалії. Показовим GNN-рішенням для поточкових даних NIDS є Anomal-E [12], яке поєднує графові векторні подання (embeddings) з класичними детекторами аномалій (зокрема Isolation Forest). Воно підтверджує цінність графової структури, але потребує GPU-навчання.

У [1] виявлення бокового руху формалізовано як задачу темпорального передбачення ребер. Попри перевагу над класичними підходами, такі моделі потребують значних корпоративних ресурсів, що підкреслює актуальність інтерпретованих ресурсно-економічних графових альтернатив.

Найближчим за постановкою задачі є дослідження [10], присвячене виявленню бокового руху на наборі LANL – сітчастій корпоративній топології з понад 12 тис. хостів та інтенсивним внутрішнім трафіком. На відміну від позитивних результатів [10], у цій роботі на зіркоподібній топології CICIDS2017 отримано суттєво нижчу якість виявлення, що узгоджується з гіпотезою про топологічну залежність ефективності локальних графових сигналів. Пряме порівняння результатів є обмеженим через відмінності у структурі даних та схемах побудови часових графів.

Розвиваються також некеровані GNN-підходи на атрибутіваних і темпоральних графах [3]. Ці методи трансдуктивні, потребують GPU-навчання та іншої інфраструктури, ніж промисловий стек Neo4j GDS. Пряме порівняння виходить за межі цієї роботи.

Проведений аналіз виокремлює прогалини, які заповнює ця робота. Негативні результати графових детекторів публікуються рідко і без атрибуції причин відмови (парадигма, класифікатор, ознаковий простір чи протокол оцінювання), а контрольні дизайни, що розділяють ці чинники [11], систематично до графових детекторів аномалій не застосовувалися; вплив топології на дискримінативність локальних графових сигналів обговорюється якісно, а позитивні результати отримано на сітчастих топологіях [10]. Інтерпретовані графові альтернативи на промисловому стеку відсутні в українській дискусії військово-орієнтованих IDS [5], [13], де увага зосереджена на поточкових класифікаторах, зокрема на наборі CICIDS2017 [13]; композитні метрики з кількох незалежних графових сигналів висвітлені переважно теоретично, без експериментальної валідації меж застосування.

Мета статті. Метою статті є діагностичне оцінювання меж застосовності локальних інтерпретованих графових сигналів як самостійного детектора мережевих аномалій із некерованою побудовою ознак на наборі CICIDS2017 та порівняти їхню ефективність із потокоорієнтованими базовими методами. Для досягнення мети визначено такі завдання:

- 1) побудувати формальну модель темпорального мультиграфа мережевого трафіку та композитної метрики аномальності $A(v, t)$;
- 2) розробити й програмно реалізувати запропонований метод із використанням Java 25, Neo4j та бібліотеки Graph Data Science (GDS), а також потокоорієнтований базовий детектор засобами PostgreSQL;
- 3) провести експериментальну перевірку на наборі CICIDS2017 та локалізувати причини деградації якості виявлення;



4) обґрунтувати межі застосування методу та сформулювати напрями подальших досліджень.

Робота має характер діагностичного дослідження на наборі CICIDS2017 (Mon+Tue+Wed, $\Delta t = 5$ хв) з використанням підходу [11]. Основними внесками є контрольний дизайн для локалізації причин деградації графових детекторів аномалій та методологічний шаблон їх аудиту.

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

Мережевий трафік представляється темпоральним мультиграфом:

$$G = (V, E, T, \phi_V, \phi_E, \tau, \alpha), \quad (1)$$

де V – скінченна множина вузлів, що представляють хости та сервіси; $E \subseteq V \times V \times \mathbb{N}$ – множина спрямованих темпоральних ребер, у якій третій компонент є ідентифікатором екземпляра взаємодії та забезпечує розрізнення паралельних ребер між однаковими парами вузлів; $T \subseteq \mathbb{R}_{\geq 0}$ – часова вісь; $\phi_V: V \rightarrow \{Host, Service\}$ та $\phi_E: E \rightarrow \{CONNECTS_TO, HOSTS\}$ – функції типізації вузлів і ребер відповідно; $\tau: E \rightarrow \mathcal{I}$ – функція, що ставить у відповідність ребру інтервал його активності $\tau(e) = [t_{start}(e), t_{end}(e))$, де $\mathcal{I}$ – множина допустимих часових інтервалів; $\alpha: E \rightarrow \mathcal{A}$ – функція атрибутів ребер (протокол, кількість агрегованих потоків, обсяг переданих даних). У подальшому аналізі використовуються лише ребра типу CONNECTS_TO, що з'єднують вузли типу Host; типи Service і HOSTS передбачені схемою для розширення моделі на рівень сервісів і до побудови ознак не залучаються. Збереження часової структури є принциповим, оскільки агрегація взаємодій у статичний граф призводить до втрати інформації про послідовність і часовий контекст мережевої активності.

Темпоральні вікна

Для дискретизації неперервного потоку подій вводиться аналітичне вікно тривалості $\Delta t > 0$. Множина ребер вікна:

$$W(t, \Delta t) = \{e \in E: t_{start}(e) \in [t - \Delta t, t)\}. \quad (2)$$

На основі $W(t, \Delta t)$ будується індукований підграф G_t , що представляє стан мережі у момент t . Додатково вводиться базове вікно $\Delta t_b \gg \Delta t$ (типово $\Delta t_b = 24$ год для $\Delta t = 5$ хв) над верифікованими «чистими» проміжками без зафіксованих інцидентів – для обчислення референсних статистик. Крок просування $\delta t \leq \Delta t$ визначає частоту оновлення вікна.

Компоненти композитної метрики

Для кожного хоста v на кожному вікні t обчислюються три показники. Компонент α_1 – відхилення центральності. Як міра використовується betweenness centrality (центральність за посередництвом) [7]:

$$BC_t(v) = \sum_{s \neq v \neq u} \frac{\sigma_{su}(v)}{\sigma_{su}}, \quad s, u \in V_t, \quad (3)$$

де σ_{su} – кількість найкоротших шляхів між s і u у графі вікна G_t , а $\sigma_{su}(v)$ – кількість таких шляхів, що проходять через v . Ця міра виявляє вузли-мости – природний сигнал для вузлів-плацдармів у боковому русі [1]. На основі базового вікна обчислюються μ_v^{BC} та σ_v^{BC} , а компонент визначається як стандартизоване відхилення:

$$\alpha_1(v, t) = \frac{|BC_t(v) - \mu_v^{BC}|}{\sigma_v^{BC} + \varepsilon}, \quad (4)$$

де $\varepsilon = 10^{-6}$ – регуляризація для запобігання діленню на нуль.

Компонент α_2 – зміна спільноти. Нехай для кожного моменту t' задано функцію community($\cdot, t'$): $V_{t'} \rightarrow \mathbb{N}$, що повертає ідентифікатор спільноти вузла у графі $G_{t'}$ за алгоритмом Louvain [6]. Зміна реєструється як бінарна подія, зважена силою приналежності до попередньої спільноти; для скорочення позначень уведемо $C_{t'}(v) \equiv \text{community}(v, t')$; так, $C: V_t \rightarrow \mathbb{N}$ – функція приналежності у поточному вікні.

$$\alpha_2(v, t) = \mathbb{1}[C_t(v) \neq C_{t-\delta t}(v)] \cdot \rho(v, t - \delta t), \quad (5)$$



де $\rho(v, t')$ – частка інцидентних v ребер вікна t' , другий кінець яких належить тій самій спільноті, що й v . Зважування через ρ відрізняє природний шум кластеризації периферійних вузлів ($\rho \approx 0$) від структурної трансформації центральних для спільноти вузлів ($\rho \rightarrow 1$).

Компонент α_3 характеризує зміну локальної околиці. Нехай $N_k^t(v) = \{u \in V_t : \text{dist}_{G_t}(v, u) \leq k\} \setminus \{v\}$ – k -стрибова околиця вузла v . У дослідженні прийнято $k = 2$. Ступінь трансформації околиці оцінюється за допомогою Жаккардового розходження (Jaccard drift):

$$\alpha_3(v, t) = 1 - \frac{|N_k^t(v) \cap N_k^{t-\delta t}(v)|}{|N_k^t(v) \cup N_k^{t-\delta t}(v)|}. \quad (6)$$

Композитна метрика та правило рішення

Три компоненти об'єднуються зваженою сумою:

$$A(v, t) = w_1 \tilde{\alpha}_1(v, t) + w_2 \alpha_2(v, t) + w_3 \alpha_3(v, t), \quad \sum_{i=1}^3 w_i = 1, \quad w_i \geq 0, \quad (7)$$

де $\tilde{\alpha}_1 = \sigma(\alpha_1 - \theta_1)$ – сигмоїдально нормалізоване значення в $[0,1]$ (компоненти α_2, α_3 уже визначені у $[0,1]$ за побудовою). Параметр θ_1 підбирається виключно на сегменті Monday (легітимний трафік, без доступу до тренувальної частини Tue+Wed) так, щоб мода $\tilde{\alpha}_1$ на цьому базовому розподілі припадала на проміжок $[0,1; 0,3]$. Цільовий діапазон обрано апіорно [3], [4]; інваріантність результатів до схеми нормалізації перевірено експериментально.

Лінійну форму композиції у (7) обрано з міркувань інтерпретовності (кожен коефіцієнт w_i відображає відносну вагу компонента), можливості систематичного пошуку на сітці у симплексі ваг і сумісності з базовим профілем. Альтернативи (максимум, геометрична середня, рангове об'єднання) не розглядаються як основні: перші посилюють артефакт інвертованої семантики α_3 , рангові не мають нативної підтримки у GDS; керовані ваги перевірено окремим контрольним експериментом.

Правило класифікації вузла v у момент t :

$$D(v, t) = \mathbf{1}[A(v, t) > \theta_A], \quad D = 1 - \text{аномальний}, \quad D = 0 - \text{легітимний}. \quad (8)$$

Параметри моделі: $\Delta t = \delta t = 5$ хв (пошук на сітці $\{2, 5, 10, 15, 30\}$ хв), $\Delta t_b = 24$ год, $\Delta t_a = 60$ с, $k = 2$, $\epsilon = 10^{-6}$, розмір вибірки наближеної центральності 1000; ваги w_1, w_2, w_3 і поріг θ_A калібруються на тренувальному сегменті максимізацією F_1 , зсув θ_1 визначається за модою $\tilde{\alpha}_1$ на базовому трафіку.

МЕТОДИКА ДОСЛІДЖЕННЯ

Архітектура обчислювального конвеєра та реалізація

Метод реалізовано як послідовність чотирьох етапів, що забезпечують послідовне оновлення та аналіз графа у Neo4j.

Етап 1. Завантаження та перетворення сирого потоку CSV-записів на темпоральний мультиграф (1) з агрегацією потоків між парою вузлів за одним протоколом у часові кошики тривалості $\Delta t_a = 60$ с.

Етап 2. Формування базового профілю передбачає обчислення $\mu_v^{BC}, \sigma_v^{BC}$ для кожного хоста за послідовністю вікон тривалості Δt у межах базового вікна Δt_b на верифікованому легітимному трафіку. Центральність обчислюється у наближеному режимі з вибіркою (значення параметрів наведено вище).

Етап 3. Виявлення аномалій у ковзному вікні здійснюється за алгоритмом, наведеним нижче.

Етап 4. На етапі оцінювання формується композиція в $A(v, t)$ за (7) і рішення (8).

Кошик агрегації $\Delta t_a = 60$ с обрано як компроміс між збереженням сигналу та обчислювальною здійсненністю: зменшення до 10–30 с збільшило б кількість супер-ребер у 2–6 разів, тоді як обране значення дає коефіцієнт стиснення $5,56\times$ (таблиця 1).

Систему реалізовано на Java 25 (Spring Boot 3.3); графова СКБД – Neo4j Community з бібліотекою Graph Data Science (GDS), реляційна – PostgreSQL 16. Стан між вікнами зберігається у властивостях вузлів графа, тому сервіси є безстановими; фільтрація вікна підтримується індексом часу початку ребра. Усі компоненти реалізовано з фіксованим зерном генератора випадкових чисел; вихідний код доступний у публічному репозиторії <https://github.com/yevheniiredziuk/anomaly-graph-ids>. Експерименти виконано на робочій станції з процесором Apple M1 Pro (10 ядер) і 16 ГБ оперативної пам'яті.

Алгоритм виявлення у ковзному вікні

У кожному вікні виконується така послідовність дій: проєкція підграфа вікна за (2); обчислення наближеної центральності за посередництвом (3) і компонента α_1 за (4) з базовими статистиками;



виявлення спільнот алгоритмом Louvain та обчислення α_2 за (5) з попереднім станом вузла; побудова 2-стрижкових околиць і обчислення α_3 за (6); композиція (7) і рішення (8). Ідентифікатор спільноти, частка внутрішньоспільнотних ребер і околиця зберігаються у властивостях вузла для наступного вікна; вікно просувається на δt .

На першому вікні після фази базового профілю попередній стан вузлів не визначено. Тому для α_2 приймається нульове значення, а α_3 за (6) дорівнює одиниці для всіх вузлів з непорожньою околицею; α_1 обчислюється звичайним чином за базовими статистиками. Для хостів, відсутніх у базовому вікні, статистики μ_v^{BC} та σ_v^{BC} не визначені, і $\tilde{\alpha}_1$ приймається рівним нулю. Проекція графа звільняється наприкінці кожної ітерації.

Асимптотична складність

Для одного вікна з $n = |V_t|$ вузлами та $m = |W(t, \Delta t)|$ ребрами:

$$O_{\text{window}} = O(k \cdot (n + m)) + O(m \log n) + O(n \cdot \bar{d}^2 + n_{\text{hub}} \cdot d_{\text{max}}^2), \quad (9)$$

наближена ВС, $k=1000$ Louvain 2-стрижкові околиці

де $\bar{d}$ – середній степінь, n_{hub} – кількість концентраторів (вузлів з високим степенем), d_{max} – максимальний степінь. У медіанному режимі (408 вузлів, 514 ребер на вікно) складність близька до $O(n \cdot m)$ і прийнятна для 5-хвилинних вікон; у найгіршому випадку концентратори з високим d_{max} кратно збільшують останній доданок.

Набір даних, протокол оцінювання та порівнювані методи

Для експериментального оцінювання використано набір даних CICIDS2017 [8], зокрема CSV-файли згенерованих розмічених потоків (GeneratedLabelledFlows). До аналізу включено три дні. Понеділок (Mon) використано для формування базового профілю, а вівторок (Tue) і середу (Wed) – для експериментального оцінювання. Четвер і п'ятницю не включено, оскільки атаки Web та Infiltration не відповідають постановці задачі на рівні хостів у часових вікнах.

Таблиця 1

Статистика попереднього оброблення

Етап	Mon	Tue	Wed	Всього
Сирі потоки	529 918	445 909	692 703	1 668 530
Після усунення дублів + NaN/Inf	529 884	445 905	692 685	1 668 474
Після агрегації ($\Delta t_a = 60$ с)	124 128	102 560	102 388	329 076

Легітимний трафік становить 84 % потоків; серед атак домінує DoS Hulk (13,85 %), решта шести класів сумарно займають близько 2 %. Понеділковий трафік використано для базового профілю, трафік Tue+Wed поділено за часом навпіл на тренувальний (калібрування порогів і ваг) і тестовий сегменти.

Еталонна розмітка задана на рівні пари «хост × вікно»: хост вважається аномальним у вікні, якщо хоча б один його потік має мітку атаки. Протокол victim-tagging позначає позитивною пару, в якій хост є джерелом або призначенням атакуючого потоку; протокол source-only розглядає лише пари, де хост є джерелом, і позначає позитивними тільки атакуювальників (таблиця 2). Останній рядок таблиці 2 відповідає розмітці source-only, застосованій до тієї самої множини 76 646 пар, що й victim-tagging (56 позитивних, базова частота 0,073 %); ця спільна множина використовується у таблиці 3 для порівняння ранжувальної здатності сигналів, щоб ізолювати вплив семантики міток від впливу складу вибірки. Значення F_1 , отримані за різними протоколами, не є безпосередньо зіставними.

Таблиця 2

Оцінні множини за протоколами розмітки

Протокол	Множина	Пар «хост × вікно»	Позитивних	Базова частота, %
victim-tagging	повна Tue+Wed	76 646	112	0,146
victim-tagging	тестова половина	38 459	70	0,182
source-only	повна Tue+Wed	50 727	56	0,110
source-only	тестова половина	25 688	35	0,136
source-only на спільній множині	повна Tue+Wed	76 646	56	0,073



Оцінна множина має зіркоподібну топологію, типову для корпоративних мереж: у медіанному вікні 408 вузлів і 514 ребер, 83,3 % вузлів – листкові (ступінь 1), діаметр графа 4, ступінь концентраторів-серверів сягає 467 у межах вікна.

Порівнюються три методи. Метод 1 – досліджуваний графовий детектор (7)–(8) у двох конфігураціях: $\text{pure-}\alpha_1$, де $w = (1, 0, 0)$, $\theta_A = 0,225$, та композитній де $w = (0,5, 0,4, 0,1)$ і $\theta_A = 0,6$. Метод 2 – Isolation Forest (IF) на 21 потоковій ознаці, агрегований до пари «хост $\times$ вікно». Метод 3 – три експертні правила, реалізовані у PostgreSQL, для виявлення сканування портів, перебір паролів та DoS-флуд. Порогові значення правил налаштовано відповідно до характеристик атак у CICIDS2017.

Мітки класів використовуються лише для калібрування порогів і ваг на тренувальному сегменті, а не для формування ознакового простору. Метод 3 ґрунтується на експертних правилах, тому основним контрольним методом обрано метод 2, тоді як метод 3 використано як додатковий орієнтир.

Якість бінарного виявлення оцінюється за точністю P, повнотою R і F_1 -мірою при калібрувальному порозі, ранжувальна здатність – за показниками PR-AUC і ROC-AUC. Довірчі інтервали (CI) 95 % отримано стратифікованим бутстрапом з 1000 повторень; відмінності між рівнями помилок перевірено за критерієм Мак-Немара.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Якість виявлення

Метод 3 перевершує графовий за повнотою на шести з семи класів атак (на DoS Hulk і GoldenEye – 1,000 проти 0,000; макроусереднена повнота 0,644 проти 0,083; через малу кількість позитивних пар у класах покласові значення є якісними тенденціями). За узагальненими метриками на повній множині Tue+Wed (victim-tagging) найкраща конфігурація методу 1 ($\text{pure-}\alpha_1$) дає $F_1 = 0,0135$ ($P = 0,0069$, $R = 0,3393$) проти $F_1 = 0,0797$ у методу 3 ($P = 0,0431$, $R = 0,5357$); метод 2 на тестовій половині у протоколі source-only досягає $F_1 = 0,2000$ ($P = 0,1132$, $R = 0,8571$). ROC-AUC методу 1 становить 0,624 (95 % CI [0,569; 0,679]) проти 0,759 [0,710; 0,808] у методу 3 і 0,987 [0,976; 0,994] у методу 2; PR-AUC методу 1 (0,0033) лише у 2,3 раза вищий за рівень випадкового детектора при базовій частоті 0,146 %, тоді як у методу 3 він дорівнює 0,0370. Помірна ранжувальна здатність не переходить у придатну точність, що видно з кривих «точність–повнота» на рис. 1.

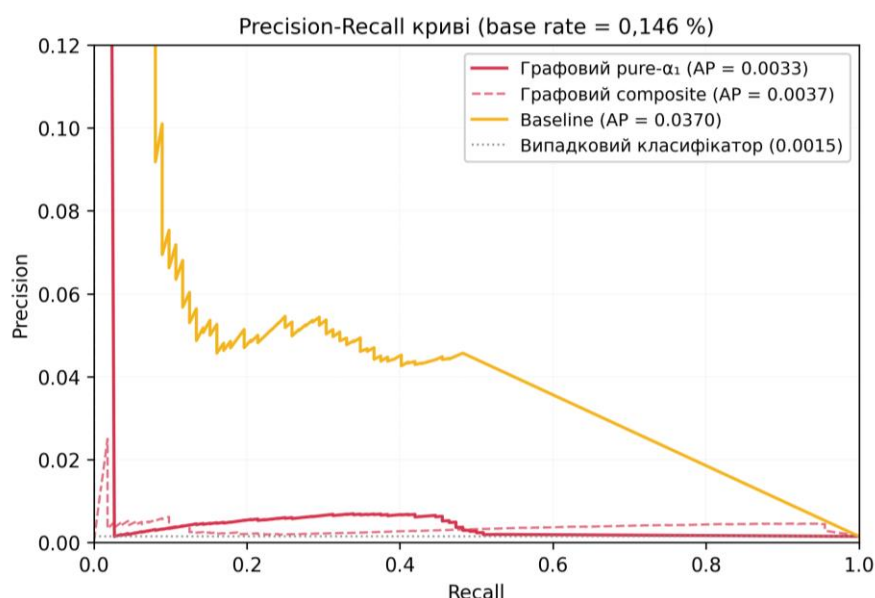


Рис. 1. Криві «точність–повнота» (базова частота позитивного класу 0,146 %)

Діагностика: абляція, ранжування та контрольні експерименти

Пошук ваг виконано на тренувальному сегменті у два рівні. Груба сітка з 8 наборів ваг (вершини, середини ребер і центроїд симплексу) визначила опорну композитну конфігурацію, дрібна сітка уточнила пороги (37 значень з кроком 0,025). Ані серед цих 296 комбінацій, ані у щільному пошуку по симплексу (231 набір ваг, крок 0,05) з перевіркою поза вибіркою не знайдено композиції, яка перевершила б однокомпонентний детектор $\text{pure-}\alpha_1$: композиції з вищим F_1 на повній множині (до 0,0196) вироджені й



при калібруванні на тренувальній половині дають на тестовій $F_1 = 0,0103$ проти 0,0132. Оптимум $\Delta t = 5$ хв стійкий на сітці $\{2, 5, 10, 15, 30\}$ хв, а заміна сигмоїдальної нормалізації α_1 на min-max або z-score змінює F_1 не більш як на 0,001.

Абляційний аналіз показує, що α_1 – єдиний компонент, який робить внесок у F_1 : його виключення знижує F_1 на 16 %, тоді як виключення α_3 результату не змінює, а окремо взяті α_2 і α_3 поступаються α_1 у 1,4 та 4,6 рази.

За ранжувальною здатністю ситуація інша (таблиця 3). У протоколі source-only ROC-AUC компонента $\tilde{\alpha}_1$ дорівнює 0,339 – центральність антикорелює з роллю атакувальника; знакокоригований $1 - \alpha_3$ і α_2 ранжують добре (0,836 і 0,839), проте їх PR-AUC перевищує випадковий рівень лише у 4–7 разів, тоді як об'ємна ознака weighted degree – у 244 рази. Отже, структурні сигнали розрізняють класи, але за базової частоти 0,146 % цього недостатньо для придатної точності навіть за оракульного порога.

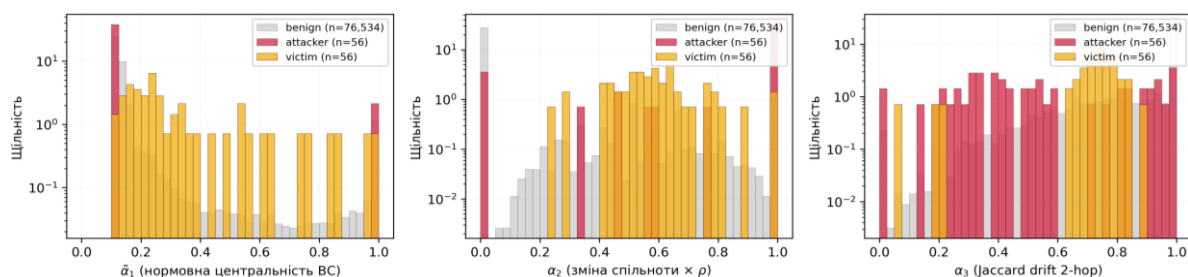
Таблиця 3

Ранжувальна здатність окремих сигналів

Сигнал	Тип	ROC-AUC victim	ROC-AUC source	PR-AUC source
$\tilde{\alpha}_1$ (відхилення ВС)	структурний	0,624	0,339	0,00080
α_2 (зміна спільноти)	структурний	0,802	0,839	0,00283
$1 - \alpha_3$ (знакокориговане)	структурний	0,817	0,836	0,00541
degree	структурний	0,841	0,706	0,00135
weighted degree	об'ємний	0,970	0,957	0,17836
Випадковий детектор	—	0,500	0,500	0,00073

Механізм відмови пояснюють розподіли компонентів для легітимних хостів, атакувальників і жертв (рис. 2). Компонент $\tilde{\alpha}_1$ ранжує жертв вище за атакувальників: медіана для атакувальників дорівнює 0,119 (сигмоїда при нульовій центральності), для жертв – 0,241, тому звітне F_1 у протоколі victim-tagging є наслідком включення жертв до позитивного класу. Компонент α_2 реагує на концентратори: алгоритм Louvain [6] об'єднує листкові вузли в одну велику спільноту, і зміна спільноти для них – шум кластеризації. Компонент α_3 має інвертовану семантику (ROC-AUC 0,18): «нова околиця» характерна для легітимних хостів зі спорадичною активністю, тоді як атакувальники повторно звертаються до тих самих жертв. Оскільки компоненти реагують на різні класи подій, зважена сума усереднює сигнал і деградує порівняно з $\text{pure-}\alpha_1$.

Контрольний дизайн розділяє чотири можливі причини низької якості: класифікатор, семантику розмітки, природу ознак і їх кількість. Кероване навчання ваг на тих самих трьох ознаках логістичною регресією (збалансовані класи, навчання на тренувальному сегменті) не покращує результату. На тестовій половині $F_1 = 0,0063$ проти 0,0132 у $\text{pure-}\alpha_1$ (CI [0,0084; 0,0186]), а від'ємний коефіцієнт при α_3 незалежно підтверджує його інвертовану семантику. Критерій Мак-Немара ($p < 10^{-4}$) фіксує лише відмінність у структурі помилок; жоден режим компромісу «точність–повнота» не виводить детектори на цих ознаках за межу $F_1 \approx 0,02$.

Рис. 2. Розподіли компонентів $\tilde{\alpha}_1$, α_2 , α_3 для легітимних хостів, атакувальників та жертв

Якщо центральність виявляє саме атакувальників, перехід від victim-tagging до source-only мав би підвищити F_1 . Результат протилежний: F_1 детектора $\text{pure-}\alpha_1$ падає у 12 разів (з 0,0135 до 0,0011), композитна конфігурація дає 0,0031; α_2 ранжує атакувальників добре ($R = 0,804$), але через дисбаланс класів дає $F_1 = 0,0062$.



Один і той самий некерований класифікатор IF застосовано до різних ознакових просторів у спільному протоколі source-only на вирівняній тестовій вибірці (25 688 пар, 35 позитивних). На трьох графових ознаках IF дає $F_1 = 0,0039$ – на рівні pure- α_1 , отже проблема не в класифікаторі. На трьох потокових ознаках (кількість потоків, обсяг переданих байтів, кількість віддалених портів) – $F_1 = 0,1506$, у 38,6 раза вище, що ізолює природу ознак від їх кількості; 5 і 21 ознака дають 0,1579 і 0,2000; 95 % СІ для потокових просторів ([0,1053; 0,1966], [0,1106; 0,2085], [0,1706; 0,2311]) не перекриваються з СІ для графових ([0,0000; 0,0154] у логістичній регресії), а критерій Мак-Немара підтверджує систематичність переваги ($p < 10^{-4}$). Елементарна структурна ознака – степінь вузла – дає $F_1 = 0,0059$, тоді як пара «степінь + зважений степінь» – 0,1069; оскільки зважений степінь дорівнює кількості потоків хоста, дискримінативність забезпечує об'ємна складова, а не структура зв'язків. Усі перевірені структурні сигнали (α_1 та α_2 окремо, трійка ($\tilde{\alpha}_1, \alpha_2, \alpha_3$) як ознаковий простір IF і степінь вузла) у протоколі source-only не перевищують $F_1 = 0,0062$.

Обчислювальна продуктивність графового конвеєра є додатковим обмеженням: пропускна здатність становить близько 0,044 вікна/с проти 44,75 у потокового (p99 латентності наближеної центральності – 156 мс на вікно), і розрив на три порядки є бар'єром для режиму реального часу.

Обмеження результатів

Експеримент проведено виключно на CICIDS2017, який навіть після очищення [9] зберігає до 5 % шуму в мітках. Узагальнення потребує верифікації на CSE-CIC-IDS2018, UNSW-NB15 і STU-13. Протокол victim-tagging включає жертв до позитивного класу, що зміщує оцінку для α_1 ; правило «один хост у вікні – одне рішення» не враховує, що хост може бути одночасно атакувальником і жертвою. Кошик $\Delta t_a = 60$ с не валідувався пошуком на сітці (інваріантність до $\Delta t_a \in \{10, 30, 60\}$ с обґрунтовано лише якісно), стабільність спільнот Louvain між вікнами кількісно не оцінювалась, а відсутність інкрементного обчислення центральності є бар'єром для потокового розгортання. Простір структурних ознак неповний, оскільки коефіцієнт кластеризації, темпоральні мотиви та динаміка ребер не тестувалися, тому висновок обмежений чотирма перевіреними сигналами: відхиленням центральності, зміною спільноти, розходженням околиці та ступенем вузла. Порівняння з GNN-методами не виконано через принципово іншу інфраструктуру. Поправку Бонферроні до конфігурацій сітки не застосовано, оскільки пошук виконувався лише на тренувальному сегменті.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Наукова новизна роботи полягає в тому, що запропоновано діагностичний шаблон для аудиту графових детекторів аномалій, який поєднує декомпозицію еталонної розмітки, контрольний дизайн «класифікатор × ознаковий простір» та розрізнення чотирьох категорій причин низької результативності; за результатами проведеного огляду, таке поєднання раніше до графових детекторів аномалій не застосовувалося. Його застосування до локальних графових ознак на темпоральних мультиграфах дало змогу встановити умови, за яких такі ознаки не забезпечують операційно прийнятної точності.

Основні результати дослідження полягають у такому.

1. Експериментально встановлено обмежену дискримінативну здатність структурних ознак у досліджуваній постановці задачі. За спільного протоколу оцінювання детектор pure- α_1 продемонстрував $F_1 = 0,0021$, тоді як IF на потокових ознаках забезпечив $F_1 = 0,2000$. Контрольні експерименти засвідчили, що виявлена розбіжність зумовлена насамперед природою структурних сигналів, а не вибором класифікатора чи кількістю ознак. Водночас окремі структурні ознаки (зміна спільноти, степінь вузла) забезпечили ранжування хостів, що перевищувало випадковий рівень (ROC-AUC до 0,84). Отже, ранжувальна здатність ознаки не гарантує її операційної придатності як самостійного детектора за низької базової частоти позитивного класу.

2. Визначено сукупність структурних і семантичних чинників, що обмежують ефективність графового виявлення аномалій: переважання листкових вузлів у досліджуваній топології, зміщена семантика розмітки victim-tagging для ознак α_1 та інвертований характер сигналу α_3 , корекція якого за знаком не забезпечує працездатності детектора.

3. Встановлено межі застосовності отриманих результатів. Висновки стосуються мережевої топології з переважанням листкових робочих станцій і фіксованого кошика агрегації потоків $\Delta t_a = 60$ с. Їх перенесення на сітчасті топології, сценарії бокового руху зі щільними підграфами та гібридні графово-потоківі ознакові простори потребує окремої експериментальної перевірки. Для досліджуваного класу корпоративних мереж результати обґрунтовують доцільність використання потокових ознак як основи виявлення аномалій; роль графових ознак як додаткового інструмента залишається відкритим питанням.



Подальші дослідження доцільно зосередити на перевірці топологічної залежності ефективності досліджених ознак на мережах з іншою структурою, зокрема на сітчастих топологіях (CTU-13, CSE-CIC-IDS2018, UNSW-NB15). Окремим напрямом є роздільний аналіз вихідних і вхідних взаємодій для точнішого розмежування поведінки атакувальника та жертви. Оскільки лінійна композиція та логістична регресія не дали переваги над однокомпонентним детектором, доцільно перевірити нелінійні моделі та інші некеровані детектори, а також розширити структурний ознаковий простір коефіцієнтом кластеризації, темпоральними мотивами та динамікою ребер. Перспективним є поєднання локальних графових ознак із потоковими характеристиками та оцінювання ефективності такого гібридного ознакового простору.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. King, I. J., & Huang, H. H. (2023). Euler: Detecting network lateral movement via scalable temporal link prediction. *ACM Transactions on Privacy and Security*, 26(3), Article 35. <https://doi.org/10.1145/3588771>
2. Chen, Z., Tang, D., Li, Q., Liu, Z., Zhang, J., & Zhang, Y. (2024). A survey on graph neural networks for intrusion detection systems: Methods, trends and challenges. *Computers & Security*, 141, Article 103821. <https://doi.org/10.1016/j.cose.2024.103821>
3. Ma, X., Wu, J., Xue, S., Yang, J., Zhou, C., Sheng, Q. Z., Xiong, H., & Akoglu, L. (2023). A comprehensive survey on graph anomaly detection with deep learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(12), 12012–12038. <https://doi.org/10.1109/TKDE.2021.3118815>
4. Akoglu, L., Tong, H., & Koutra, D. (2015). Graph-based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery*, 29(3), 626–688. <https://doi.org/10.1007/s10618-014-0365-y>
5. Ільєнко, А., Ільєнко, С., Яковенко, О., Галич, Є., & Павленко, В. (2024). Перспективи інтеграції штучного інтелекту в системи кібербезпеки. *Кібербезпека: освіта, наука, техніка*, 1(25), 318–329. <https://doi.org/10.28925/2663-4023.2024.25.318329>
6. Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), Article P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
7. Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2), 163–177. <https://doi.org/10.1080/0022250X.2001.9990249>
8. Sharafaldin, I., Habibi Lashkari, A., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. In *Proceedings of the 4th International Conference on Information Systems Security and Privacy* (pp. 108–116). SCITEPRESS. <https://doi.org/10.5220/0006639801080116>
9. Engelen, G., Rimmer, V., & Joosen, W. (2021). Troubleshooting an intrusion detection dataset: The CICIDS2017 case study. In *2021 IEEE Security and Privacy Workshops* (pp. 7–12). IEEE. <https://doi.org/10.1109/SPW53761.2021.00009>
10. Bowman, B., Laprade, C., Ji, Y., & Huang, H. H. (2020). Detecting lateral movement in enterprise computer networks with unsupervised graph AI. In *23rd International Symposium on Research in Attacks, Intrusions and Defenses* (pp. 257–268). USENIX Association. <https://www.usenix.org/conference/raid2020/presentation/bowman>
11. Arp, D., Quiring, E., Pendlebury, F., Warnecke, A., Pierazzi, F., Wressnegger, C., Cavallaro, L., & Rieck, K. (2022). Dos and don'ts of machine learning in computer security. In *Proceedings of the 31st USENIX Security Symposium* (pp. 3971–3988). USENIX Association. <https://www.usenix.org/conference/usenixsecurity22/presentation/arp>
12. Caville, E., Lo, W. W., Layeghy, S., & Portmann, M. (2022). Anomal-E: A self-supervised network intrusion detection system based on graph neural networks. *Knowledge-Based Systems*, 258, Article 110030. <https://doi.org/10.1016/j.knosys.2022.110030>
13. Мальцева, І. Р., Черниш, Ю. О., & Процюк, Ю. О. (2024). Аналіз алгоритмів раннього виявлення кібератак на мережі з використанням машинного навчання. *Системи і технології зв'язку, інформатизації та кібербезпеки*, (6), 105–115. <https://doi.org/10.58254/viti.6.2024.08.105>

**Yevhenii Redziuk**

PhD in Technical Sciences,

Head of the Department of Computer Science and Intelligent Technologies

Military Institute of Telecommunications and Informatization named after Heroes of Kruty, Kyiv, Ukraine

ORCID: 0000-0001-5592-5121

yevhenii.redziuk@viti.edu.ua

**FACTORIAL DIAGNOSTICS OF THE DISCRIMINATIVE POWER OF LOCAL GRAPH FEATURES
IN ANOMALY DETECTION ON TEMPORAL MULTIGRAPHS**

Abstract. Anomaly detection on temporal interaction graphs is a data-mining task in which flow-based methods capture only a limited structural context of interactions between nodes. The aim of this work is to assess the discriminative power of local interpretable graph features as a stand-alone detector of network anomalies under label-free feature construction. A composite anomaly score is proposed based on three graph components: deviation of betweenness centrality from a baseline profile, Louvain community change and Jaccard drift of the local neighbourhood. The experimental evaluation is carried out on the CICIDS2017 dataset using a controlled classifier $\times$ feature-space design and a comparison with flow-based baselines. The method is implemented on the Neo4j Graph Data Science and Java stack for the analysis of a temporal multigraph of network traffic. The results show that the studied graph features retain ranking ability (ROC-AUC up to 0.84) but do not provide operationally acceptable precision as a stand-alone detector: under the common evaluation protocol, $F_1 = 0.0021$ versus $F_1 = 0.2000$ for the flow-based baseline. Control experiments localise the source of degradation in the feature space rather than in the classifier, the number of features or the unsupervised nature of feature construction. The scientific novelty of the work lies in establishing the limits of the discriminative power of local graph features in the studied setting and in developing a diagnostic template for auditing graph-based anomaly detectors that combines ground-truth decomposition, controlled comparison of classifiers and feature spaces, and analysis of the structural and semantic causes of low performance. The results justify the use of flow-based features as the basis of anomaly detection in the studied class of enterprise networks; the generalisation of the conclusions to mesh topologies requires separate experimental verification.

Keywords: data mining; graph analytics; anomaly detection; temporal multigraph; graph features; community detection; unsupervised learning; intrusion detection.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. King, I. J., & Huang, H. H. (2023). Euler: Detecting network lateral movement via scalable temporal link prediction. *ACM Transactions on Privacy and Security*, 26(3), Article 35. <https://doi.org/10.1145/3588771>
2. Chen, Z., Tang, D., Li, Q., Liu, Z., Zhang, J., & Zhang, Y. (2024). A survey on graph neural networks for intrusion detection systems: Methods, trends and challenges. *Computers & Security*, 141, Article 103821. <https://doi.org/10.1016/j.cose.2024.103821>
3. Ma, X., Wu, J., Xue, S., Yang, J., Zhou, C., Sheng, Q. Z., Xiong, H., & Akoglu, L. (2023). A comprehensive survey on graph anomaly detection with deep learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(12), 12012–12038. <https://doi.org/10.1109/TKDE.2021.3118815>
4. Akoglu, L., Tong, H., & Koutra, D. (2015). Graph-based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery*, 29(3), 626–688. <https://doi.org/10.1007/s10618-014-0365-y>
5. Iliencko, A., Iliencko, S., Yakovenko, O., Halych, Ye., & Pavlenko, V. (2024). Prospects for the integration of artificial intelligence into cybersecurity systems. *Cybersecurity: Education, Science, Technique*, 1(25), 318–329. <https://doi.org/10.28925/2663-4023.2024.25.318329> [in Ukrainian]
6. Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), Article P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
7. Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2), 163–177. <https://doi.org/10.1080/0022250X.2001.9990249>
8. Sharafaldin, I., Habibi Lashkari, A., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. In *Proceedings of the 4th International Conference on Information Systems Security and Privacy* (pp. 108–116). SCITEPRESS. <https://doi.org/10.5220/0006639801080116>



9. Engelen, G., Rimmer, V., & Joosen, W. (2021). Troubleshooting an intrusion detection dataset: The CICIDS2017 case study. In 2021 IEEE Security and Privacy Workshops (pp. 7–12). IEEE. <https://doi.org/10.1109/SPW53761.2021.00009>
10. Bowman, B., Laprade, C., Ji, Y., & Huang, H. H. (2020). Detecting lateral movement in enterprise computer networks with unsupervised graph AI. In 23rd International Symposium on Research in Attacks, Intrusions and Defenses (pp. 257–268). USENIX Association. <https://www.usenix.org/conference/raid2020/presentation/bowman>
11. Arp, D., Quiring, E., Pendlebury, F., Warnecke, A., Pierazzi, F., Wressnegger, C., Cavallaro, L., & Rieck, K. (2022). Dos and don'ts of machine learning in computer security. In Proceedings of the 31st USENIX Security Symposium (pp. 3971–3988). USENIX Association. <https://www.usenix.org/conference/usenixsecurity22/presentation/arp>
12. Caville, E., Lo, W. W., Layeghy, S., & Portmann, M. (2022). Anomal-E: A self-supervised network intrusion detection system based on graph neural networks. Knowledge-Based Systems, 258, Article 110030. <https://doi.org/10.1016/j.knosys.2022.110030>
13. Maltseva, I. R., Chernysh, Yu. O., & Protsiuk, Yu. O. (2024). Analysis of algorithms for early detection of cyberattacks on networks using machine learning. Communication, Informatization and Cybersecurity Systems and Technologies, (6), 105–115. <https://doi.org/10.58254/viti.6.2024.08.105> [in Ukrainian]

Отримано редакцією журналу / Received: 28.05.26

Прорецензовано / Revised: 16.06.26

Схвалено до друку / Accepted: 26.10.26



This work is licensed under Creative Commons Attribution-noncommercial-sharealike 4.0 International License.